







```
FT Domain 1..15
FT /note= "measles virus F protein helper T cell epitope"
FT Domain 16..25
FT /note= "LHRH hapten"
XX
XX
PN WO9425060-A1.
XX
XX 10-NOV-1994.
XX
XX 28-APR-1994; 94WO-US04832.
XX
XX 27-APR-1993; 93US-0057166.
XX 14-APR-1994; 94US-0229275.
XX
XX (LADD/) LADD A E.
XX (WANG/) WANG C Y.
XX (ZAMB/) ZAMB T.
XX
XX Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
XX
XX Immunogenic luteinising hormone releasing hormone peptide(s) -
XX that suppress LHRH activity in males and females
XX
XX Claim 8; Page 84; 213pp; English.
XX
XX Synthetic immunogenic peptides are provided in which a universal immune
XX stimulator is linked to a peptide or protein hapten containing B cell
XX and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX potent immune responses to the coupled peptide or protein. The
XX stimulator consists of (A) a promiscuous helper T cell epitope (Th)
XX which elicits an immune response to the coupled peptide in members of
XX a heterogeneous population expressing diverse HLA phenotypes, and (B)
XX an adjuvant peptide sequence from the invasive protein of Yersinia.
XX Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
XX invasive and Th domains and between the immune stimulator and haptens
XX components. When the haptens is LHRH, then optionally the invasive domain
XX can be omitted from the immune stimulator component.
XX The present sequence represents an LHRH-containing, invasive-free
XX immunogenic peptide as above which can be used as a potent vaccine for
XX treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
XX prostatic carcinoma, testicular carcinoma, endometriosis, benign
XX uterine tumours, recurrent functional ovarian cysts, (severe)
XX premenstrual syndrome or oestrogen-dependent breast cancer, or for
XX induction of infertility.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 25 AA;
XX
XX Query Match 73.1%; Score 117; DB 15; Length 25;
XX Best Local Similarity 85.7%; Pred. No. 1.2e-10;
XX Matches 24; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
XX
OY 3 LSEIKGVIVHRLEGVEGSLHWSYGLRP 30
XX |||||||
DB 1 LSEIKGVIVHRLEGVE-----HWSYGLRP 24
XX
RESULT 4
AAR62708
ID AAR62708 standard; peptide; 42 AA.
XX
AC AAR62708;
XX
XX 25-MAR-2003 (updated)
DT 10-SEP-1995 (first entry)
XX
XX LHRH-containing immunogenic peptide.
XX
XX Helper T cell epitope; universal immune stimulator; invasive; hapten;
XX vaccine; LHRH; luteinising hormone releasing hormone; prostate;
XX androgen-dependent carcinoma; antitumour; infertility;
KW
```

```
KW measles virus F protein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Domain 1..15
FT /note= "measles virus F protein helper T cell epitope"
FT Domain 16..30
FT /note= "measles virus F protein helper T cell epitope"
FT Domain 33..42
FT /note= "LHRH hapten"
XX
XX WO9425060-A1.
XX
XX 10-NOV-1994.
XX
XX 28-APR-1994; 94WO-US04832.
XX
XX 27-APR-1993; 93US-0057166.
XX 14-APR-1994; 94US-0229275.
XX
XX (LADD/) LADD A E.
XX (WANG/) WANG C Y.
XX (ZAMB/) ZAMB T.
XX
XX Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
XX
XX Immunogenic luteinising hormone releasing hormone peptide(s) -
XX that suppress LHRH activity in males and females
XX
XX Claim 8; Page 86; 213pp; English.
XX
XX Synthetic immunogenic peptides are provided in which a universal immune
XX stimulator is linked to a peptide or protein hapten containing B cell
XX and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX potent immune responses to the coupled peptide or protein. The
XX stimulator consists of (A) a promiscuous helper T cell epitope (Th)
XX which elicits an immune response to the coupled peptide in members of
XX a heterogeneous population expressing diverse HLA phenotypes, and (B)
XX an adjuvant peptide sequence from the invasive protein of Yersinia.
XX Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
XX invasive and Th domains and between the immune stimulator and haptens
XX components. When the haptens is LHRH, then optionally the invasive domain
XX can be omitted from the immune stimulator component.
XX The present sequence represents an LHRH-containing, invasive-free
XX immunogenic peptide as above which can be used as a potent vaccine for
XX treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
XX uterine tumours, recurrent functional ovarian cysts, (severe)
XX premenstrual syndrome or oestrogen-dependent breast cancer, or for
XX induction of infertility.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 42 AA;
XX
XX Query Match 73.1%; Score 117; DB 15; Length 42;
XX Best Local Similarity 82.8%; Pred. No. 2.3e-10;
XX Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
XX
OY 2 LLEIKGVIVHRLEGVEGSLHWSYGLRP 30
XX :|||||
DB 15 VLSEIKGVIVHRLEGVGGE--HWSYGLRP 41
XX
RESULT 5
AAR62707
ID AAR62707 standard; peptide; 27 AA.
XX
AC AAR62707;
XX
XX 25-MAR-2003 (updated)
DT
```

PD 15-NOV-2001.  
 XX  
 XX  
 XX 04-MAY-2001; 2001WO-US14363.  
 XX  
 XX 05-MAY-2000; 2000US-202328P.  
 XX  
 XX (APHT-) APHTON CORP.  
 XX  
 XX Grimes S, Michaeli D, Stevens VC;  
 XX  
 XX WPI; 2002-049440/06.  
 DR  
 XX  
 XX Novel synthetic immunogen for inducing immune response against  
 XX gonadotropin releasing hormone, comprises fusion peptide having  
 XX promiscuous helper T-cell peptide epitope and immunomimic peptide  
 XX epitope or its analogue -  
 XX  
 XX Claim 11; Page 7; 43pp; English.  
 PS  
 XX The invention relates to a synthetic immunogen for inducing specific  
 XX antibodies against gonadotropin releasing hormone (GnRH) also known  
 XX as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 XX peptide which comprises a promiscuous helper T-cell peptide epitope and  
 XX immunomimic peptide epitope or its analogue. The synthetic  
 XX immunogen is useful inducing an immune response against GnRH in an  
 XX animal subject, and as such is useful as a contraceptive and in the  
 XX treatment of diseases such as cancer (of the breast, uterus and other  
 XX gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 XX hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 XX high and specific anti-GnRH antibody titres. The present sequence  
 XX is a synthetic immunogen of the invention.  
 XX  
 SQ Sequence 31 AA;  
 Query Match 99.4%; Score 159; DB 23; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-17;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30  
 |||||  
 DB 1 KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30  
 |||||  
 RESULT 2  
 AAU11428  
 ID AAU11428 standard; peptide; 47 AA.  
 XX  
 AC AAU11428;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Synthetic immunogen peptide 9.  
 XX  
 KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Plasmodium falciparum.  
 OS Chimeric - Mammalia.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 1  
 FT /label= OTHER  
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
 FT 1..10  
 FT Peptide  
 FT /note= "Gonadotropin releasing hormone epitope  
 FT (1..10 aa)"  
 FT Peptide  
 FT 11..16  
 FT /note= "Spacer peptide"  
 FT 17..34

FT Peptide /note= "Malaria CSP protein (288-302 aa)"  
 35..38  
 FT /note= "Spacer peptide"  
 39..47  
 FT /note= "Gonadotropin releasing hormone epitope  
 (2-10 aa)"  
 FT Modified-site 47  
 FT /note= "Amidated glycine or glycylamide"  
 XX  
 XX WO200185763-A2.  
 PN  
 XX  
 XX 15-NOV-2001.  
 PD  
 XX  
 XX 04-MAY-2001; 2001WO-US14363.  
 XX  
 XX 05-MAY-2000; 2000US-202328P.  
 PR  
 XX  
 XX (APHT-) APHTON CORP.  
 XX  
 XX Grimes S, Michaeli D, Stevens VC;  
 XX  
 XX WPI; 2002-049440/06.  
 DR  
 XX Novel synthetic immunogen for inducing immune response against  
 XX gonadotropin releasing hormone, comprises fusion peptide having  
 XX promiscuous helper T-cell peptide epitope and immunomimic peptide  
 XX epitope or its analogue -  
 XX  
 XX Claim 11; Page 11; 43pp; English.  
 PS  
 XX The invention relates to a synthetic immunogen for inducing specific  
 XX antibodies against gonadotropin releasing hormone (GnRH) also known  
 XX as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 XX peptide which comprises a promiscuous helper T-cell peptide epitope and  
 XX immunomimic peptide epitope or its analogue. The synthetic  
 XX immunogen is useful inducing an immune response against GnRH in an  
 XX animal subject, and as such is useful as a contraceptive and in the  
 XX treatment of diseases such as cancer (of the breast, uterus and other  
 XX gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 XX hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 XX high and specific anti-GnRH antibody titres. The present sequence  
 XX is a synthetic immunogen of the invention.  
 XX  
 SQ Sequence 47 AA;  
 Query Match 99.4%; Score 159; DB 23; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-16;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30  
 |||||  
 DB 17 KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 46  
 |||||  
 RESULT 3  
 AAAR62705  
 ID AAAR62705 standard; peptide; 25 AA.  
 XX  
 AC AAAR62705;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 10-SEP-1995 (first entry)  
 XX  
 XX LHRH-containing immunogenic peptide.  
 DE  
 XX  
 KW Helper T cell epitope; universal immune stimulator; invasin; haptens;  
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;  
 KW androgen-dependent carcinoma; antitumour; infertility;  
 KW measles virus F protein.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:58:36 ; Search time 35.036 Seconds  
(without alignments)  
140.442 Million cell updates/sec

Title: US-09-848-834A-9  
Perfect score: 160  
Sequence: 1 KLLSEIKGVIVHREGEVPSLHWSYGLRPX 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	159	99.4	31	AAU11420	Synthetic immunogen
2	159	99.4	47	AAU11428	Synthetic immunogen
3	117	73.1	25	AARG2705	LHRH-containing im
4	117	73.1	42	AARG2708	LHRH-containing im
5	116	72.5	27	AARG2707	LHRH-containing im
6	116	72.5	27	AA191156	Peptide immunogen
7	116	72.5	27	AA191156	Peptide immunogen
8	116	72.5	45	AARG2721	LHRH-containing im
9	113	70.6	27	AA191163	Modified MVF Th ep

10	110	68.8	31	21	AA191175	Modified MVF Th ep
11	108	67.5	27	21	AA191161	Modified MVF Th ep
12	108	67.5	27	21	AA191167	Modified MVF Th ep
13	108	67.5	45	21	AA191165	Modified MVF Th ep
14	108	67.5	45	21	AA191165	Peptide immunogen
15	107	66.9	31	21	AA191179	Modified MVF Th ep
16	106	66.2	28	21	AA191158	Modified MVF Th ep
17	105	65.6	31	21	AA191173	Modified MVF Th ep
18	105	65.6	31	21	AA191173	Peptide immunogen
19	105	65.6	47	21	AA191180	Inv epitope/modifi
20	105	65.6	47	21	AA191180	Peptide immunogen
21	105	65.6	49	21	AA191177	Modified MVF Th ep
22	103	64.4	27	21	AA191170	Modified MVF Th ep
23	103	64.4	27	21	AA191170	Peptide immunogen
24	103	64.4	35	21	AA191242	Modified MVF Th ep
25	100	62.5	28	15	AARG2726	LHRH-containing im
26	100	62.5	46	15	AARG2728	LHRH-containing im
27	100	62.5	47	21	AA191183	Inv epitope/modifi
28	100	62.5	47	21	AA191183	Peptide immunogen
29	99	61.9	28	21	AA191159	Modified MVF Th ep
30	99	61.9	75	24	ABP72236	Frizzled putative
31	94	58.8	28	21	AA191157	Modified MVF Th ep
32	89	55.6	27	16	AA191157	SSALI THLHRH. Sy
33	87	54.4	34	17	AAU05619	MVFRH-GG-mige2s.
34	86	53.8	34	23	AAU11424	Synthetic immunogen
35	84	52.5	27	21	AA191171	Modified MVF Th ep
36	84	52.5	27	21	AA191171	Peptide immunogen
37	84	52.5	45	21	AA191172	Modified MVF Th ep
38	84	52.5	45	21	AA191172	Peptide immunogen
39	81	50.6	31	21	AA191184	Modified MVF Th ep
40	81	50.6	31	21	AA191184	Peptide immunogen
41	79	49.4	27	21	AA191164	Modified MVF Th ep
42	79	49.4	27	21	AA191164	Modified MVF Th ep
43	79	49.4	27	21	AA191168	Modified MVF Th ep
44	79	49.4	27	21	AA191169	Modified MVF Th ep
45	79	49.4	27	21	AA191169	Peptide immunogen
			34	23	AAU11421	Synthetic immunogen

ALIGNMENTS

RESULT 1  
AAU11420  
ID AAU11420 standard; peptide; 31 AA.  
XX  
AC AAU11420;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Synthetic immunogen peptide 1.  
XX  
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
KW luteinising hormone releasing hormone; LHRH; contraceptive;  
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
XX  
OS Chimeric - Measles virus.  
OS Chimeric - Mammalia.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Peptide /note= "Measles virus fusion protein F epitope"  
FT Peptide 19..22  
FT Peptide /note= "Spacer peptide"  
FT Peptide 23..31  
FT Modified-site 31  
FT /note= "Gonadotrophin releasing hormone epitope"  
XX  
PN WO200185763-A2.  
XX

```
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-223-711-11

Query Match      45.6%; Score 73; DB 15; Length 40;
Best Local Similarity 94.1%; Pred. No. 0.0016;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 KLLSEIKGVIVHRLEGV 17
        |||||
Db      23 KLLSLIKGVIVHRLEGV 39
```

```
Search completed: September 24, 2003, 17:33:06
Job time : 15.5429 secs
```

; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Measles virus  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(16)  
; OTHER INFORMATION: Amino acid sequence 288-302 of the measles  
; OTHER INFORMATION: virus fusion protein, F  
US-09-848-834A-8

Query Match 48.1%; Score 77; DB 9; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGVE 18  
| | | | | | | | | | | | | | | |  
DB 1 LSEIKGVIVHRLGVE 16

RESULT 13  
US-09-848-834A-11  
; Sequence 11, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 830-844 of the  
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to am  
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone  
; NAME/KEY: MOD.RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Amidated-glutamine  
; NAME/KEY: MOD.RES  
; LOCATION: (28)  
; OTHER INFORMATION: Amidated-glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(15)  
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus Toxoid Precursor  
; OTHER INFORMATION: (tentoxylisin)  
; NAME/KEY: PEPTIDE  
; LOCATION: (16)..(19)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (20)..(28)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-11

Query Match 45.6%; Score 73; DB 9; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GPSLHWSYGLRP 30  
| | | | | | | | | | | | | | | | | | | | | |  
DB 16 GPSLHWSYGLRP 27

RESULT 14  
US-09-848-834A-12  
; Sequence 12, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 378-398 of  
; OTHER INFORMATION: Plasmodium falciparum CSP protein linked by a spacer to amino  
; OTHER INFORMATION: id sequence 2-10 of the GnRH hormone  
; NAME/KEY: MOD.RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Amidated aspartic acid  
; NAME/KEY: MOD.RES  
; LOCATION: (33)..(33)  
; OTHER INFORMATION: Amidated glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(20)  
; OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria  
; OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite  
; OTHER INFORMATION: (CSP) protein  
; NAME/KEY: PEPTIDE  
; LOCATION: (21)..(24)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (25)..(33)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-12

Query Match 45.6%; Score 73; DB 9; Length 33;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GPSLHWSYGLRP 30  
| | | | | | | | | | | | | | | | | | | | | |  
DB 21 GPSLHWSYGLRP 32

RESULT 15  
US-10-223-711-11  
; Sequence 11, Application US/10223711  
; Publication No. US20030113344A1  
; GENERAL INFORMATION:  
; APPLICANT: Bakaletz, Lauren O.  
; APPLICANT: Kaumaya, Pravin T.P.  
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides  
; FILE REFERENCE: 18525/04058  
; CURRENT APPLICATION NUMBER: US/10/223,711  
; CURRENT FILING DATE: 2002-08-19  
; PRIOR APPLICATION NUMBER: 09/148,711  
; PRIOR FILING DATE: 1998-09-04  
; PRIOR APPLICATION NUMBER: 08/460,502  
; PRIOR FILING DATE: 1995-06-02  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11

; NAME/KEY: MOD\_RES  
; LOCATION: (50)..(50)  
; OTHER INFORMATION: Amidated glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(37)  
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent  
; NAME/KEY: PEPTIDE  
; LOCATION: (38)..(41)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (42)..(50)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-18

Query Match 49.4%; Score 79; DB 9; Length 50;  
Best Local Similarity 92.9%; Pred. No. 0.00028;  
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 17 VEGPSLHWSYGLRP 30  
:|||||  
Db 36 LEGPSLHWSYGLRP 49

## RESULT 9

US-10-223-711-10  
; Sequence 10, Application US/10223711  
; Publication No. US20030113344A1  
; GENERAL INFORMATION:  
; APPLICANT: Bakaletz, Lauren O.  
; APPLICANT: Kaumaya, Pravin T.P.  
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides  
; FILE REFERENCE: 18525/04058  
; CURRENT APPLICATION NUMBER: US/10/223,711  
; CURRENT FILING DATE: 2002-08-19  
; PRIOR APPLICATION NUMBER: 09/148,711  
; PRIOR FILING DATE: 1998-09-04  
; PRIOR APPLICATION NUMBER: 08/460,502  
; PRIOR FILING DATE: 1995-06-02  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Artificial.  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-223-711-10

Query Match 48.8%; Score 78; DB 15; Length 40;  
Best Local Similarity 94.4%; Pred. No. 0.00031;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLSEIKGVIVHRLGVE 18  
|||||  
Db 23 KLLSLIKGVIVHRLGVE 40

## RESULT 10

US-09-847-102A-31  
; Sequence 31, Application US/09847102A  
; Publication No. US20030044409A1  
; GENERAL INFORMATION:  
; APPLICANT: University of California  
; APPLICANT: Carson, Dennis A.  
; APPLICANT: Corr, Maripat  
; APPLICANT: Rhee, Chae-Seo

; APPLICANT: Lorenzo, Leoni M.  
; APPLICANT: Malini, Sen  
; TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS  
; FILE REFERENCE: 22000-20629.00  
; CURRENT APPLICATION NUMBER: US/09/847,102A  
; CURRENT FILING DATE: 2001-05-01  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 31  
; LENGTH: 75  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: PFZD2-MMVF  
US-09-847-102A-31

Query Match 48.8%; Score 78; DB 11; Length 75;  
Best Local Similarity 94.4%; Pred. No. 0.00062;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLSEIKGVIVHRLGVE 18  
|||||  
Db 58 KLLSLIKGVIVHRLGVE 75

## RESULT 11

US-10-285-976-229  
; Sequence 229, Application US/10285976  
; Publication No. US20030165500A1  
; GENERAL INFORMATION:  
; APPLICANT: Rhee, Chae-Seo  
; APPLICANT: Malini, Sen  
; APPLICANT: Wu, Christina  
; APPLICANT: Leoni, Lorenzo M.  
; APPLICANT: Corr, Maripat  
; APPLICANT: Carson, Dennis A.  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy  
; TITLE OF INVENTION: in Head and Neck Squamous Cell Carcinomas  
; FILE REFERENCE: 023070-130320US  
; CURRENT APPLICATION NUMBER: US/10/285,976  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: US 60/287,995  
; PRIOR FILING DATE: 2001-05-01  
; PRIOR APPLICATION NUMBER: WO PCT/US02/13802  
; PRIOR FILING DATE: 2002-05-01  
; NUMBER OF SEQ ID NOS: 232  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 229  
; LENGTH: 75  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:PFZD2-MMVF  
; OTHER INFORMATION: measles virus fusion (MVF) epitope fused to  
; OTHER INFORMATION: frizzled domain  
US-10-285-976-229

Query Match 48.8%; Score 78; DB 12; Length 75;  
Best Local Similarity 94.4%; Pred. No. 0.00062;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLSEIKGVIVHRLGVE 18  
|||||  
Db 58 KLLSLIKGVIVHRLGVE 75

## RESULT 12

US-09-848-834A-8  
; Sequence 8, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:

```
; PRIOR APPLICATION NUMBER: US 60/287,995
; PRIOR FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: WO PCT/US02/13802
; PRIOR FILING DATE: 2002-05-01
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 231
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:PMVP-2D2
; OTHER INFORMATION: measles virus fusion (MVF) epitope fused to
; OTHER INFORMATION: frizzled domain
US-10-285-976-231

Query Match          61.9%; Score 99; DB 12; Length 75;
Best Local Similarity 95.5%; Pred. No. 5.6e-07;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 KLLSEIKGVIVHRLGVEGPSL 22
Db      2 KLLSLIKGVIVHRLGVEGPSL 23

RESULT 6
US-09-848-834A-13
; Sequence 13, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the G
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of t
; OTHER INFORMATION: he Measles virus fusion protein,
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(18)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(34)
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles
; OTHER INFORMATION: virus fusion protein, F
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
US-09-848-834A-13

Query Match          53.8%; Score 86; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KLLSEIKGVIVHRLGVE 18
Db      17 KLLSEIKGVIVHRLGVE 34

RESULT 7
US-09-848-834A-10
```

```
; Sequence 10, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated phenylalanine
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(21)
; OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (22)..(25)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (26)..(34)
; OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (34)..(34)
; OTHER INFORMATION: Amidated glycine or glycinamide
US-09-848-834A-10

Query Match          49.4%; Score 79; DB 9; Length 34;
Best Local Similarity 92.9%; Pred. No. 0.00018;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      17 VEGPSLHWSYGLRP 30
Db      20 LEGPSLHWSYGLRP 33

RESULT 8
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of h
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the
; OTHER INFORMATION: anus toxoid precursor (Tentoxylisin) protein linked by a sp
; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
```

OTHER INFORMATION: Amidated glycine or glycineamide  
US-09-848-834A-9

Query Match 99.4%; Score 159; DB 9; Length 31;  
Best Local Similarity 100.0%; Pred. No. 4.3e-16;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30  
|||||  
Db 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30

## RESULT 2

US-09-848-834A-17  
Sequence 17, Application US/09848834A  
Patent No. US20020076416A1

GENERAL INFORMATION:  
APPLICANT: Aption Corporation  
TITLE OF INVENTION: Chimeric Peptide Immunogens  
FILE REFERENCE: 1102865-0047  
CURRENT APPLICATION NUMBER: US/09/848,834A  
CURRENT FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: 60/202,328  
PRIOR FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 17  
LENGTH: 47  
TYPE: PRT  
ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gr  
OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of  
OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid seq  
OTHER INFORMATION: uence 2-10 of the GnRH hormone  
NAME/KEY: MOD.RES  
LOCATION: (1)..(1)  
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
NAME/KEY: MOD.RES  
LOCATION: (47)..(47)  
OTHER INFORMATION: Amidated-glycine or glycineamide  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(10)  
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
NAME/KEY: PEPTIDE  
LOCATION: (11)..(18)  
OTHER INFORMATION: Spacer peptide  
NAME/KEY: PEPTIDE  
LOCATION: (19)..(34)  
OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,  
NAME/KEY: PEPTIDE  
LOCATION: (35)..(38)  
OTHER INFORMATION: Spacer peptide  
NAME/KEY: PEPTIDE  
LOCATION: (39)..(47)  
OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone

## US-09-848-834A-17

Query Match 99.4%; Score 159; DB 9; Length 47;  
Best Local Similarity 100.0%; Pred. No. 6.9e-16;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 30  
|||||  
Db 17 KLLSEIKGVIVHRLEGVEGPSLHWSYGLRP 46

## RESULT 3

US-10-076-674-9  
Sequence 9, Application US/10076674  
Publication No. US20030165478A1

GENERAL INFORMATION:  
APPLICANT: Sokoll, Kenneth K.

TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System  
FILE REFERENCE: Immunogen Delivery System  
CURRENT APPLICATION NUMBER: US/10/076,674  
CURRENT FILING DATE: 2002-04-23  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 9  
LENGTH: 45  
TYPE: PRT  
ORGANISM: Human  
US-10-076-674-9

Query Match 72.5%; Score 116; DB 12; Length 45;  
Best Local Similarity 85.7%; Pred. No. 1.1e-09;  
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30  
|||||  
Db 19 LSEIKGVIVHRLEGVGGE--HWSYGLRP 44

## RESULT 4

US-09-847-102A-33  
Sequence 33, Application US/09847102A  
Publication No. US20030044409A1

GENERAL INFORMATION:  
APPLICANT: University of California  
APPLICANT: Carson, Dennis A.  
APPLICANT: Corr, Maripat  
APPLICANT: Rhee, Chae-Seo  
APPLICANT: Lorenzo, Leoni M.  
APPLICANT: Malini, Sen  
TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR  
FILE REFERENCE: 22000-20629, 00  
CURRENT APPLICATION NUMBER: US/09/847,102A  
CURRENT FILING DATE: 2001-05-01  
NUMBER OF SEQ ID NOS: 138  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 33  
LENGTH: 75  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: PMMVF-FZD2  
US-09-847-102A-33

Query Match 61.9%; Score 99; DB 11; Length 75;  
Best Local Similarity 95.5%; Pred. No. 5.6e-07;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLSEIKGVIVHRLEGVEGPSL 22  
|||||  
Db 2 KLLSLIKGVIVHRLEGVEGPSL 23

## RESULT 5

US-10-285-976-231  
Sequence 231, Application US/10285976  
Publication No. US20030165500A1

GENERAL INFORMATION:  
APPLICANT: Rhee, Chae-Seo  
APPLICANT: Malini, Sen  
APPLICANT: Wu, Christina  
APPLICANT: Leoni, Lorenzo M.  
APPLICANT: Corr, Maripat  
APPLICANT: Carson, Dennis A.  
TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy  
FILE REFERENCE: 023070-130320US  
CURRENT APPLICATION NUMBER: US/10/285,976  
CURRENT FILING DATE: 2002-11-01

Query Match 61.9%; Score 99; DB 11; Length 75;  
Best Local Similarity 95.5%; Pred. No. 5.6e-07;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:57:10 ; Search time 15.5429 Seconds  
(without alignments)  
301.778 Million cell updates/sec

Title: US-09-848-834A-9  
Perfect score: 160  
Sequence: 1 KLLSEIKGVIVHRLGVSGPSLHWSYGLRPX 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues  
Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				Published_Applications_AA:*	
1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*				
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*				
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*				
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*				
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*				
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*				
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*				
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*				
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*				
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*				
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*				
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*				
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*				
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*				
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*				
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*				
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*				
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	159	99.4	31	9	US-09-848-834A-9
2	159	99.4	47	9	US-09-848-834A-17
3	116	72.5	45	12	US-10-076-674-9
4	99	61.9	75	11	US-09-847-102A-33
5	99	61.9	75	12	US-10-285-976-231
6	86	53.8	34	9	US-09-848-834A-13
7	79	49.4	34	9	US-09-848-834A-10
8	79	49.4	50	9	US-09-848-834A-18
9	78	48.8	40	15	US-10-223-711-10
10	78	48.8	75	11	US-09-847-102A-31
11	78	48.8	75	12	US-10-285-976-229
12	77	48.1	16	9	US-09-848-834A-8
13	73	45.6	28	9	US-09-848-834A-11
14	73	45.6	33	9	US-09-848-834A-12
15	73	45.6	40	15	US-10-223-711-11

16	73	45.6	46	9	US-09-848-834A-19	Sequence 19, Appl
17	73	45.6	51	9	US-09-848-834A-20	Sequence 20, Appl
18	72	45.0	15	11	US-09-747-802-16	Sequence 16, Appl
19	72	45.0	15	11	US-09-747-802-30	Sequence 30, Appl
20	72	45.0	15	11	US-09-865-294-8	Sequence 8, Appl
21	72	45.0	15	11	US-09-865-294-22	Sequence 22, Appl
22	72	45.0	15	12	US-10-261-446-20	Sequence 20, Appl
23	72	45.0	550	10	US-09-873-233A-18	Sequence 18, Appl
24	72	45.0	550	10	US-09-873-233A-20	Sequence 20, Appl
25	69	43.1	15	11	US-09-747-802-37	Sequence 37, Appl
26	69	43.1	15	11	US-09-865-294-29	Sequence 29, Appl
27	69	43.1	16	15	US-10-223-711-3	Sequence 3, Appl
28	69	43.1	19	11	US-09-747-802-48	Sequence 48, Appl
29	69	43.1	19	11	US-09-865-294-40	Sequence 40, Appl
30	66	41.2	19	11	US-09-747-802-51	Sequence 51, Appl
31	66	41.2	19	11	US-09-865-294-43	Sequence 43, Appl
32	65	40.6	662	11	US-09-951-061A-141	Sequence 141, App
33	64	40.0	15	11	US-09-747-802-35	Sequence 35, Appl
34	64	40.0	15	11	US-09-747-802-38	Sequence 38, Appl
35	64	40.0	15	11	US-09-747-802-40	Sequence 40, Appl
36	64	40.0	15	11	US-09-747-802-42	Sequence 42, Appl
37	64	40.0	15	11	US-09-865-294-27	Sequence 27, Appl
38	64	40.0	15	11	US-09-865-294-30	Sequence 30, Appl
39	64	40.0	15	11	US-09-865-294-32	Sequence 32, Appl
40	64	40.0	15	11	US-09-865-294-34	Sequence 34, Appl
41	64	40.0	19	11	US-09-747-802-49	Sequence 49, Appl
42	64	40.0	19	11	US-09-747-802-55	Sequence 55, Appl
43	64	40.0	19	11	US-09-865-294-38	Sequence 38, Appl
44	64	40.0	19	11	US-09-865-294-41	Sequence 41, Appl
45	64	40.0	19	11	US-09-865-294-47	Sequence 47, Appl

ALIGNMENTS

US-09-848-834A-9

Sequence 9, Application US/09848834A

Patent No. US20020076416A1

GENERAL INFORMATION:

APPLICANT: Aphton Corporation

TITLE OF INVENTION: Chimeric Peptide Immunogens

FILE REFERENCE: 1102865-0047

CURRENT APPLICATION NUMBER: US/09/848,834A

CURRENT FILING DATE: 2001-05-04

PRIOR APPLICATION NUMBER: 60/202,328

PRIOR FILING DATE: 2000-05-05

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn version 3.0

SEQ ID NO 9

LENGTH: 31

TYPE: PPT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Chimeric peptide made up of amino acid sequence 288-302 of the asels virus fusion protein, F linked by a spacer peptide to an acid sequence 2-10 of the GnRH hormone

OTHER INFORMATION: asels virus fusion protein, F linked by a spacer peptide to an acid sequence 2-10 of the GnRH hormone

NAME/KEY: MOD\_RES

LOCATION: (1)..(1)

OTHER INFORMATION: Amidated Lysine

NAME/KEY: PEPTIDE

LOCATION: (1)..(15)

OTHER INFORMATION: Peptide corresponds to the amino acid sequences 288-302 of the measles virus fusion protein, F

NAME/KEY: PEPTIDE

LOCATION: (19)..(22)

OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE

LOCATION: (23)..(31)

OTHER INFORMATION: Peptide corresponds to amino acid sequences 2-10 of the human GnRH hormone

NAME/KEY: MOD\_RES

LOCATION: (31)..(31)



REFERENCE/DOCKET NUMBER: 1151-4157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-770-014-55

Query Match 68.8%; Score 110; DB 4; Length 31;  
Best Local Similarity 78.6%; Pred. No. 3.6e-10;  
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 3 LSEIKGVIVHRLGVGPGSLHWSYGLRP 30  
Db 3 LSEIKGVIVHRLGVGPGSLHWSYGLRP 30

Search completed: September 24, 2003, 17:00:21  
Job time : 9.28837 secs

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-770-014-43

Query Match      70.6%; Score 113; DB 4; Length 27;
Best Local Similarity 82.1%; Pred. No. 1.1e-10;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 3 LSEIKGVIVHRLGEGVPSLHWSYGLRP 30
   |||||:|||||:|||||
Db 1 LSEIKGVIVHKLEGVGGE--HWSYGLRP 26

RESULT 13
US-09-100-414B-55
; Sequence 55, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-55

Query Match      68.8%; Score 110; DB 3; Length 31;
Best Local Similarity 78.6%; Pred. No. 3.6e-10;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLGEGVPSLHWSYGLRP 30
   |||||:|||||:|||||
Db 3 LSEIKGVIVHKLEGVGGEHWSYGLRP 30

RESULT 14
US-09-303-323-55
; Sequence 55, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-55

Query Match      68.8%; Score 110; DB 3; Length 31;
Best Local Similarity 78.6%; Pred. No. 3.6e-10;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLGEGVPSLHWSYGLRP 30
   |||||:|||||:|||||
Db 3 LSEIKGVIVHKLEGVGGEHWSYGLRP 30

RESULT 15
US-09-770-014-55
; Sequence 55, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
```

```
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGEGPSLHWSYGLRP 30
    ||||| ||||| ||||| ||||| |||||
Db 19 LSEIKGVIVHRLGEGVGE--HWSYGLRP 44

RESULT 10
US-09-100-414B-43
; Sequence 43, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-43

Query Match 70.6%; Score 113; DB 3; Length 27;
Best Local Similarity 82.1%; Pred. No. 1.le-10;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGEGPSLHWSYGLRP 30
    ||||| ||||| ||||| ||||| |||||
Db 1 LSEIKGVIVHRLGEGVGE--HWSYGLRP 26

RESULT 11
US-09-303-323-43
; Sequence 43, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-43

Query Match 70.6%; Score 113; DB 3; Length 27;
Best Local Similarity 82.1%; Pred. No. 1.le-10;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGEGPSLHWSYGLRP 30
    ||||| ||||| ||||| ||||| |||||
Db 1 LSEIKGVIVHRLGEGVGE--HWSYGLRP 26

RESULT 12
US-09-770-014-43
; Sequence 43, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
```

Query Match 72.5%; Score 116; DB 2; Length 45;  
Best Local Similarity 85.7%; Pred. No. 6.7e-11;

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/229,275  
;; FILING DATE: 14-APR-1994  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/057,166  
;; FILING DATE: 27-APR-1992  
;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Marla H. Lin  
;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)415-8745  
;; TELEFAX: (516)751-6849  
;; INFORMATION FOR SEQ ID NO: 19:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 27 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-488-351A-19

Query Match 72.5%; Score 116; DB 2; Length 27;  
Best Local Similarity 85.7%; Pred. No. 3.7e-11;  
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30  
|||||  
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 5  
US-09-100-414B-36  
;; Sequence 36, Application US/09100414B  
;; Patent No. 6025468  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, Chang Yi  
;; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
;; NUMBER OF SEQUENCES: 106  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Morgan & Finnegan, L.L.P.  
;; STREET: 345 Park Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10154-0054  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC Compatible  
;; OPERATING SYSTEM: PC Windows  
;; SOFTWARE: Word 97  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/100,414B  
;; FILING DATE: 20-JUNE-1998  
;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Marla H. Lin  
;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4157  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-758-4800  
;; TELEFAX: 212-751-6849  
;; INFORMATION FOR SEQ ID NO: 36:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 27 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-100-414B-36

Query Match 72.5%; Score 116; DB 3; Length 27;

Best Local Similarity 85.7%; Pred. No. 3.7e-11;  
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;  
QY 3 LSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30  
|||||  
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 6  
US-09-303-323-36  
;; Sequence 36, Application US/09303323  
;; Patent No. 6228987  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, Chang Yi  
;; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
;; NUMBER OF SEQUENCES: 106  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Morgan & Finnegan, L.L.P.  
;; STREET: 345 Park Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10154-0054  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC Compatible  
;; OPERATING SYSTEM: PC Windows  
;; SOFTWARE: Word 97  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/303,323  
;; FILING DATE: 30-APR-1999  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 09/100,414  
;; FILING DATE: 20-JUNE-1998  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Marla H. Lin  
;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4157  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-758-4800  
;; TELEFAX: 212-751-6849  
;; INFORMATION FOR SEQ ID NO: 36:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 27 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-303-323-36

Query Match 72.5%; Score 116; DB 3; Length 27;  
Best Local Similarity 85.7%; Pred. No. 3.7e-11;  
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30  
|||||  
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 7  
US-09-770-014-36  
;; Sequence 36, Application US/09770014  
;; Patent No. 659282  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, Chang Yi  
;; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
;; NUMBER OF SEQUENCES: 106  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Morgan & Finnegan, L.L.P.  
;; STREET: 345 Park Avenue  
;; CITY: New York

```
RESULT 2
US-08-488-351A-20
; Sequence 20, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-20

Query Match 73.1%; Score 117; DB 2; Length 42;
Best Local Similarity 82.8%; Pred. No. 4.3e-11;
Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 2 LLEIKGVIVHRLEGVEGSLHWSYGLRP 30
Db 15 VLSEIKGVIVHRLEGVGGE--HWSYGLRP 41

RESULT 3
US-08-446-692-19
; Sequence 19, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-19

Query Match 72.5%; Score 116; DB 1; Length 27;
Best Local Similarity 85.7%; Pred. No. 3.7e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 3 LSEIKGVIVHRLEGVEGSLHWSYGLRP 30
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 4
US-08-488-351A-19
; Sequence 19, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-19
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:46:20 ; Search time 9.18836 Seconds  
(without alignments)  
142.750 Million cell updates/sec

Title: US-09-848-834A-9  
Perfect score: 160  
Sequence: 1 KLLSEIKGVIVHRLGVEGSLHWSYGLRXP 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues  
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA: \*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep: \*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	73.1	42	1	US-08-446-692-20
2	117	73.1	42	2	US-08-488-351A-20
3	116	72.5	27	1	US-08-446-692-19
4	116	72.5	27	2	US-08-488-351A-19
5	116	72.5	27	3	US-09-100-414B-36
6	116	72.5	27	3	US-09-303-323-36
7	116	72.5	27	4	US-09-770-014-36
8	116	72.5	45	1	US-08-446-692-33
9	116	72.5	45	2	US-08-488-351A-33
10	113	70.6	27	3	US-09-100-414B-43
11	113	70.6	27	3	US-09-303-323-43
12	113	70.6	27	4	US-09-770-014-43
13	110	68.8	31	3	US-09-100-414B-55
14	110	68.8	31	3	US-09-303-323-55
15	110	68.8	31	4	US-09-770-014-55
16	108	67.5	27	3	US-09-100-414B-41
17	108	67.5	27	3	US-09-100-414B-47
18	108	67.5	27	3	US-09-303-323-47
19	108	67.5	27	3	US-09-303-323-41
20	108	67.5	27	4	US-09-770-014-41
21	108	67.5	27	4	US-09-770-014-47
22	108	67.5	45	3	US-09-100-414B-45
23	108	67.5	45	3	US-09-303-323-45
24	108	67.5	45	4	US-09-770-014-45
25	107	66.9	31	3	US-09-100-414B-59
26	107	66.9	31	3	US-09-303-323-59
27	107	66.9	31	4	US-09-770-014-59

28	106	66.2	28	3	US-09-100-414B-38	Sequence 38, Appl
29	106	66.2	28	3	US-09-303-323-38	Sequence 38, Appl
30	106	66.2	28	4	US-09-770-014-38	Sequence 38, Appl
31	105	65.6	31	3	US-09-100-414B-53	Sequence 53, Appl
32	105	65.6	31	3	US-09-303-323-53	Sequence 53, Appl
33	105	65.6	31	4	US-09-770-014-53	Sequence 53, Appl
34	105	65.6	47	3	US-09-100-414B-60	Sequence 60, Appl
35	105	65.6	47	3	US-09-303-323-60	Sequence 60, Appl
36	105	65.6	47	4	US-09-770-014-60	Sequence 60, Appl
37	105	65.6	49	3	US-09-100-414B-57	Sequence 57, Appl
38	105	65.6	49	3	US-09-303-323-57	Sequence 57, Appl
39	105	65.6	49	4	US-09-770-014-57	Sequence 57, Appl
40	103.5	64.7	25	1	US-08-446-692-17	Sequence 17, Appl
41	103.5	64.7	25	2	US-08-488-351A-17	Sequence 17, Appl
42	103	64.4	27	3	US-09-100-414B-50	Sequence 50, Appl
43	103	64.4	27	3	US-09-303-323-50	Sequence 50, Appl
44	103	64.4	27	4	US-09-770-014-50	Sequence 50, Appl
45	103	64.4	35	3	US-09-100-414B-80	Sequence 80, Appl

ALIGNMENTS

RESULT 1  
US-08-446-692-20  
; Sequence 20, Application US/08446692  
; Patent No. 5759551  
; GENERAL INFORMATION:  
; APPLICANT: Ladd, Anna  
; APPLICANT: Wang, Chang Yi  
; APPLICANT: Zamb, Timothy  
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maria C.H. Lin  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: US  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,692  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria C.H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)415-8745  
; TELEFAX: (516)751-6849  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-446-692-20

Query Match 73.1%; Score 117; DB 1; Length 42;  
Best Local Similarity 82.8%; Pred. No. 4.3e-11;  
Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;  
QY 2 LLSKTKGVIVHRLGVEGSLHWSYGLR 30  
:|||||  
DB 15 VLSEIKGVIVHRLGVEGSLHWSYGLR 41  
:|||||





DR GO: GO:0005625; C:soluble fraction; TAS.  
 DR GO: GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.  
 DR GO: GO:0007267; P:cell-cell signaling; TAS.  
 DR GO: GO:0007275; P:development; TAS.  
 DR GO: GO:0008285; P:negative regulation of cell proliferation; TAS.  
 DR GO: GO:0007165; P:signal transduction; TAS.  
 DR InterPro: IPR002012; GNRH.  
 DR InterPro: IPR004079; GonadoliberinI.  
 DR Pfam: PF00446; GNRH; 1.  
 DR PRINTS: PR01541; GONADOLIBRNI.  
 DR PROSITE: PS00473; GNRH; 1.  
 DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 DR Placenta; Signal; Pyroglutamate carboxylic acid.  
 DR SIGNAL 1 23 PROGNADOLIBERIN I.  
 DR CHAIN 24 92 GONADOLIBERIN I.  
 DR PEPTIDE 24 33 GONADOLIBERIN I.  
 DR PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.  
 DR ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.  
 FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 FT VARIANT 16 16 W->S (IN dhsnp:6185).  
 FT SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;  
 Query Match 35.9%; Score 57.5; DB 1; Length 92;  
 Best Local Similarity 80.0%; Pred. No. 0.18;  
 Matches 12; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 QY 17 VEG-PSLHWSYGLRP 30  
 DB 18 VEGCSQHSYGLRP 32  
 RESULT 13  
 GONI\_MOUSE  
 ID GONI\_MOUSE STANDARD; PRT; 90 AA.  
 AC P13562;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Prognadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor  
 DE I].  
 DE GNRH1 OR GNRH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8706928; PubMed=3024317;  
 RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,  
 RA Phillips H.S., Nikolic K., Seeburg P.H.;  
 RA "A deletion truncating the gonadotropin-releasing hormone gene is  
 RA responsible for hypogonadism in the hpg mouse.";  
 RL Science 234:1366-1371(1986).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the GNRH family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: M14872; AAA37717.1; -.  
 DR PIR: A47578; RHMSG.  
 DR MGI: MGI:95789; Gnrh.  
 DR InterPro: IPR002012; GNRH.  
 DR InterPro: IPR004079; GonadoliberinI.  
 DR Pfam: PF00446; GNRH; 1.  
 DR PRINTS: PR01541; GONADOLIBRNI.  
 DR PROSITE: PS00473; GNRH; 1.  
 DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 DR Placenta; Signal; Pyroglutamate carboxylic acid.  
 DR SIGNAL 1 21 PROGNADOLIBERIN I.  
 DR CHAIN 22 90 GONADOLIBERIN I.  
 DR PEPTIDE 22 31 GONADOLIBERIN I.  
 DR PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.  
 DR ACT\_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.  
 FT MOD\_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).  
 FT SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;  
 Query Match 35.6%; Score 57; DB 1; Length 90;  
 Best Local Similarity 70.6%; Pred. No. 0.21;  
 Matches 12; Conservative 0; Mismatches 3; Indels 2; Gaps 1;  
 QY 14 LEGVEGSLHWSYGLRP 30  
 DB 16 LEGCS--SQHWSYGLRP 30  
 RESULT 14  
 GONI\_RAT  
 ID GONI\_RAT STANDARD; PRT; 92 AA.  
 AC P07490;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Prognadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor  
 DE I].  
 DE GNRH1 OR GNRH.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86094338; PubMed=2867548;  
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;  
 RA "Isolation of the gene and hypothalamic cDNA for the common precursor  
 RA of gonadotropin-releasing hormone and prolactin release-inhibiting  
 RA factor in human and rat.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89384661; PubMed=2476669;  
 RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;  
 RA "The rat gonadotropin-releasing hormone gene: SH locus: structure and  
 RA hypothalamic expression.";  
 RL Cell. Mol. Neurobiol. 12:447-454(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Thymus;  
 RX MEDLINE=93105480; PubMed=1468115;  
 RA Maier C.C., Marchetti B., Leboeuf R.D., Bialock J.E.;  
 RA "Thymocytes express a mRNA that is identical to hypothalamic  
 RA luteinizing hormone-releasing hormone mRNA.";  
 RL Cell. Mol. Neurobiol. 12:447-454(1992).  
 RN [4]  
 RP SEQUENCE OF 1-47 FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=87149087; PubMed=3547652;  
 RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;

```

RP SEQUENCE OF 24-33.
RX MEDLINE-72114303; PubMed-4946067;
RA Baba Y., Matsuo H., Schally A.V.;
RT "Structure of the porcine LH- and FSH-releasing hormone. II.
RT Confirmation of the proposed structure by conventional sequential
RT analyses."
RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
RN [3]
RP SYNTHESIS OF GONADOLIBERIN.
RX MEDLINE-72065376; PubMed-4942726;
RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
RT phase method."
RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
RN [4]
RP SYNTHESIS OF GONADOLIBERIN.
RX MEDLINE-72117544; PubMed-4946275;
RA Baba Y., Arimura A., Schally A.V.;
RT "On the tryptophan residue in porcine LH and FSH-releasing hormone."
RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L32864; AAA31066.1;
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 91
FT PEPTIDE 24 33
FT PEPTIDE 34 91
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;

Query Match 35.9%; Score 57.5; DB 1; Length 91;
Best Local Similarity 50.0%; Pred. No. 0.18;
Matches 15; Conservative 4; Mismatches 8; Indels 3; Gaps 1;

QY 1 KLLSEIKGVVHRLEGVGSLHWSYGLRP 30
DB 6 KLLA---GLLLLTICVGCSSQHSYGLRP 32

RESULT 12
GON1_HUMAN
ID GON1_HUMAN STANDARD; PRT; 92 AA.
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor (Contains: Gonadoliberin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
DE peptide I).
GN GnRH1 OR GnRH OR LHRH.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89366682; PubMed=2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
RT hormone gene."
RL Nucleic Acids Res. 17:6403-6403(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat."
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT SER-16.
RX MEDLINE-85012739; PubMed=6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
RT releasing hormone."
RL Nature 311:666-668(1984).
RN [4]
RP SEQUENCE OF 24-33.
RX MEDLINE-83126573; PubMed=6760865;
RA Tan L., Roussseau P.;
RT "The chemical identity of the immunoreactive LHRH-like peptide
RT biosynthesized in the human placenta."
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
RN [5]
RP VARIANT SER-16.
RX MEDLINE-99318093; PubMed=10391209;
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes."
RL Nat. Genet. 22:231-238(1999).
RN [6]
RP ERRATUM.
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
RA Lander E.S.;
RL Nat. Genet. 23:373-373(1999).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC Luterpulse or Lutrelief (Ferring Pharmaceuticals) and Relisorm
CC (Serono).
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01059; CAA25526.1;
DR EMBL; M12578; AAA35916.1;
DR EMBL; X15215; CAA33285.1;
DR F01; S05308; RHHUG.
DR Genew; HGNC:4419; GnRH1.
DR MIM; 152760;

```

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X16566; CAA34567.1; -;  
CC EMBL; X16566; CAA34568.1; ALT\_INIT.  
CC HSP; P04849; ISVF.  
CC  
CC InterPro; IPR000776; Fusion\_gly. 1.  
CC Pfam; PF00523; Fusion\_gly. 1.  
CC Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
CC SIGNAL 1 26  
CC CHAIN 27 529 FUSION GLYCOPROTEIN F0.  
CC CHAIN 27 115 PROTEIN F2.  
CC CHAIN 116 529 PROTEIN F1.  
CC TRANSMEM 116 139 POTENTIAL.  
CC DOMAIN 140 497 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 498 518 POTENTIAL.  
CC DOMAIN 519 529 CYTOPLASMIC (POTENTIAL).  
CC DISULFID 71 198 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
CC CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC SEQUENCE 529 AA; 57331 MW; AE987BC9P07E9AA9 CRC64;  
CC  
CC Query Match 40.0%; Score 64; DB 1; Length 529;  
CC Best Local Similarity 93.3%; Pred. No. 0.12;  
CC Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
CC  
CC 3 LSEIKGVIVHRLEGV 17  
CC 291 LSEIKGVIVHRLEGV 305  
CC  
CC RESULT 10  
CC VGLF\_PHODV STANDARD; PRT; 631 AA.  
CC ID VGLF\_PHODV STANDARD; PRT; 631 AA.  
CC P28886;  
CC 01-DEC-1992 (Rel. 24, Created)  
CC 01-DEC-1992 (Rel. 24, Last sequence update)  
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
CC Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
CC Fusion glycoprotein F1].  
CC GN F.  
CC Phocine distemper virus (PDV).  
CC OS Viruses; ssRNA negative-strand viruses; Mononegavirales;  
CC OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
CC OC NCBI\_TaxID=11240;  
CC [1]  
CC SEQUENCE FROM N.A.  
CC RC STRAIN=U1ster/88;  
CC RA MEDLINE=92113536; PubMed=1765768;  
CC KOevamees J., Blixenkron-Moeller M., Sharma B., Oervell C.,  
CC Norby E.;  
CC "The nucleotide sequence and deduced amino acid composition of the  
CC haemagglutinin and fusion proteins of the morbillivirus phocid  
CC distemper virus.";  
CC RT J. Gen. Virol. 72:2959-2966(1991).  
CC [2]  
CC SEQUENCE FROM N.A.  
CC RC STRAIN=U1ster/88;  
CC RA MEDLINE=92398437; PubMed=1524494;  
CC Curran M.D., Lu Y.J., Rima B.K.;  
CC "The fusion protein gene of phocine distemper virus: nucleotide and  
CC deduced amino acid sequences and a comparison of morbillivirus fusion  
CC proteins";  
CC RT Arch. Virol. 126:159-169(1992).  
CC [3]

RP SEQUENCE OF 95-631 FROM N.A.  
RC STRAIN=U1ster/88;  
RX MEDLINE=91089508; PubMed=2264246;  
RA Curran M.D., Loan D.O., Rima B.K., Kennedy S.;  
RT "Nucleotide sequence analysis of phocine distemper virus reveals its  
RT distinctness from canine distemper virus.";  
RL Vet. Rec. 127:430-431(1990).  
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
CC MEMBRANES.  
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
CC LINKED BY A DISULFIDE BOND.  
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D10371; BAA01206.1; -;  
CC PIR; A48346; A48346.  
CC PIR; JQ1368; VGN2PD.  
CC HSP; P04849; ISVF.  
CC InterPro; IPR000776; Fusion\_gly.  
CC Pfam; PF00523; Fusion\_gly. 1.  
CC Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
CC SIGNAL 1 631  
CC CHAIN ? 631 FUSION GLYCOPROTEIN F0.  
CC CHAIN ? 188 F2 PROTEIN.  
CC CHAIN 194 631 F1 PROTEIN.  
CC DISULFID 149 276 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
CC TRANSMEM 89 106 POTENTIAL.  
CC TRANSMEM 194 212 POTENTIAL.  
CC TRANSMEM 575 595 POTENTIAL.  
CC CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CONFLICT 63 I -> V (IN REF. 2).  
CC SEQUENCE 631 AA; 68873 MW; D1FC87CDD426E9B8 CRC64;  
CC  
CC Query Match 40.0%; Score 64; DB 1; Length 631;  
CC Best Local Similarity 80.0%; Pred. No. 0.14;  
CC Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
CC  
CC 3 LSEIKGVIVHRLEGV 17  
CC 369 LSEIKGVIVHRLEAV 383  
CC  
CC RESULT 11  
CC GONI\_PIG STANDARD; PRT; 91 AA.  
CC ID GONI\_PIG STANDARD; PRT; 91 AA.  
CC AC P49921;  
CC 01-OCT-1996 (Rel. 34, Created)  
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)  
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
CC Progonadolibin I precursor [Contains: Gonadolibin I (LH-RH I)  
CC (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
CC hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].  
CC GN GNRH1 OR GNRH.  
CC OS Sus scrofa (pig).  
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
CC OC NCBI\_TaxID=9823;  
CC [1]  
CC SEQUENCE FROM N.A.  
CC RC TISSUP-Hypothalamus;  
CC Weesner G.D., Matteri R.L., Becker B.A.;  
CC Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
CC [2]

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M21514; AAA47400.1; -;  
 DR PIR; A31051; VGNZK.  
 DR HSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.

FT SIGNAL 1 19 FUSION GLYCOPROTEIN F0.  
 FT CHAIN 20 546 F2 PROTEIN.  
 FT CHAIN 20 108 F1 PROTEIN.  
 FT DOMAIN 104 108 ARG-RICH (BASIC).  
 FT TRANSMEM 109 133 POTENTIAL.  
 FT TRANSMEM 484 513 POTENTIAL.  
 FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).  
 FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 546 AA; 58662 MW; 476D74DC18BCFCF CRC64;

Query Match 40.6%; Score 65; DB 1; Length 546;  
 Best Local Similarity 86.7%; Pred. No. 0.086;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGV 17  
 |||||:|||||  
 Db 284 LSEIKGVIVHRLGV 298

RESULT 8  
 VGLF\_CDVO STANDARD; PRT; 662 AA.  
 AC P12569; Q65991;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
 DE Fusion glycoprotein F1].  
 GN F.  
 OS Canine distemper virus (strain Onderstepoort) (CDV).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
 OX NCBI\_TaxID=11233;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88129050; PubMed=3433924;  
 RA Barrett T., Clarke D.K., Evans S.A., Rima B.K.;  
 RT "The nucleotide sequence of the gene encoding the F protein of canine  
 RT distemper virus: a comparison of the deduced amino acid sequence with  
 RT other paramyxoviruses.";  
 RL Virus Res. 8:373-386(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93227696; PubMed=8470428;  
 RA Wild T.F., Bernard A., Spehner D., Villevall D., Drillien R.;  
 RT "Vaccination of mice against canine distemper virus-induced  
 RT encephalitis with vaccinia virus recombinants encoding measles or  
 RT canine distemper virus antigens.";  
 RL Vaccine 11:438-444(1993).  
 CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
 CC MEMBRANES.  
 CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
 CC LINKED BY A DISULFIDE BOND.  
 CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M21849; AAA42878.1; -;  
 DR EMBL; X65509; CAA46481.1; -;  
 DR PIR; JS0321; VGNZCD.  
 DR PIR; S21382; S21382.  
 DR HSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.

FT SIGNAL 1 662 FUSION GLYCOPROTEIN F0.  
 FT CHAIN 2 224 PROTEIN F2.  
 FT CHAIN 225 662 PROTEIN F1.  
 FT TRANSMEM 606 629 POTENTIAL.  
 FT DISULFID 180 307 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 3 3 R -> K (IN REF. 2).  
 FT CONFLICT 140 140 D -> N (IN REF. 2).  
 FT CONFLICT 152 152 N -> S (IN REF. 2).  
 FT CONFLICT 171 171 I -> M (IN REF. 2).  
 FT CONFLICT 174 174 A -> V (IN REF. 2).  
 FT CONFLICT 662 662 L -> H (IN REF. 2).  
 SQ SEQUENCE 662 AA; 72970 MW; FB2C81C9797805F0 CRC64;

Query Match 40.6%; Score 65; DB 1; Length 662;  
 Best Local Similarity 86.7%; Pred. No. 0.1;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGV 17  
 |||||:|||||  
 Db 400 LSEIKGVIVHRLGV 414

RESULT 9  
 VGLF\_MEASI STANDARD; PRT; 529 AA.  
 ID VGLF\_MEASI  
 AC P26031; Q83298;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
 DE Fusion glycoprotein F1].  
 GN F.  
 OS Measles virus (strain IP-3-Ca) (Subacute sclerosing panencephalitis  
 OS virus).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
 OX NCBI\_TaxID=11237;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92263801; PubMed=15856558;  
 RA Schmidt A., Spielhofer P., Cattaneo R., Baczko K., Ter Meulen V.,  
 RA Billiet M.A.;  
 RT "Subacute sclerosing panencephalitis is typically characterized by  
 RT alterations in the fusion protein cytoplasmic domain of the  
 RT persisting measles virus.";  
 RL Virology 188:910-915(1992).  
 CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
 CC MEMBRANES.  
 CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
 CC LINKED BY A DISULFIDE BOND.  
 CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN



DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
 DE Fusion glycoprotein F1].  
 GN F.  
 OS Measles virus (strain Edmonston) (Subacute sclerosing panencephalitis  
 OS virus).  
 OS Measles virus (strain Halle) (Subacute sclerosing panencephalitis  
 OS virus).  
 OS Measles virus (strain Leningrad-16) (Subacute sclerosing panencephalitis  
 OS virus).  
 OS Measles virus (strain Edmonston-Zagreb) (Subacute sclerosing  
 OS panencephalitis virus).  
 OS Measles virus (strain Philadelphia-26) (Subacute sclerosing  
 OS panencephalitis virus), and  
 OS Measles virus (strain Edmonston B) (Subacute sclerosing panencephalitis  
 OS virus).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses;  
 OC NCBI\_TaxID=11235, 11236, 70147, 70149, 70148, 70146;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-Edmonston;  
 RC MEDLINE=87071668; PubMed=3788062;  
 RX Richardson C.D., Hull D., Greer P., Hasel K., Berkovich A.,  
 RA Englund G., Bellini W.J., Rima B., Lazzarini R.A.,  
 RA "The nucleotide sequence of the mRNA encoding the fusion protein of  
 RT measles virus (Edmonston strain): a comparison of fusion proteins  
 RT from several different paramyxoviruses.";  
 RT Virology 155:508-523(1986).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-Halle;  
 RC MEDLINE=87224816; PubMed=3585281;  
 RX Buckland R., Gerald C., Barker R., Wild T.F.;  
 RA "Fusion glycoprotein of measles virus: nucleotide sequence of the  
 RT gene and comparison with other paramyxoviruses.";  
 RL J. Gen. Virol. 68:1695-1703(1987).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-Edmonston;  
 RC MEDLINE=90085790; PubMed=2596022;  
 RX Cattaneo R., Schmid A., Spielhofer P., Kaelin K., Bacsko K.,  
 RA Meulen V., Pardowitz J., Flanagan S., Rima B.K., Udem S.A.;  
 RA "Mutated and hypermutated genes of persistent measles viruses which  
 RT caused lethal human brain diseases.";  
 RT Virology 173:415-425(1989).  
 RL [4]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-Edmonston;  
 RC MEDLINE=92263801; PubMed=1585658;  
 RX Schmid A., Spielhofer P., Cattaneo R., Bacsko K., Ter Meulen V.,  
 RA Billeter M.A.;  
 RA "Subacute sclerosing panencephalitis is typically characterized by  
 RT alterations in the fusion protein cytoplasmic domain of the  
 RT persisting measles virus.";  
 RL Virology 188:910-915(1992).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-Edmonston, Leningrad-16, and Edmonston-Zagreb;  
 RC MEDLINE=94249283; PubMed=8191786;  
 RX Rota J.S., Wang Z.D., Rota P.A., Bellini W.J.;  
 RA "Comparison of sequences of the H, F, and N coding genes of measles  
 RT virus vaccine strains.";  
 RL Virus Res. 31:317-330(1994).  
 RL [6]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-Philadelphia-26;  
 RC MEDLINE=94303181; PubMed=8030232;  
 RX Hummel K.B., Vanchiere J.A., Bellini W.J.;  
 RA "Restriction of fusion protein mRNA as a mechanism of measles virus  
 RT persistence.";

RL Virology 202:665-672(1994).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Edmonston B;  
 RA Billeter M.A.;  
 RA Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
 CC MEMBRANES.  
 CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
 CC LINKED BY A DISULFIDE BOND.  
 CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; MI4915; AAA46423.1; -  
 DR EMBL; X05597; CAA29090.1; ALT\_INIT.  
 DR EMBL; K01711; AAA75498.1; ALT\_INIT.  
 DR EMBL; K01711; AAA75499.1; -  
 DR EMBL; U03657; AAA56647.1; ALT\_INIT.  
 DR EMBL; U03659; AAA56649.1; ALT\_INIT.  
 DR EMBL; U03670; AAA56660.1; ALT\_INIT.  
 DR EMBL; U08416; AAA50501.1; ALT\_INIT.  
 DR EMBL; Z66517; CAA91367.1; ALT\_INIT.  
 DR EMBL; Z66517; CAA91368.1; -  
 DR HSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
 FT SIGNAL 1 23  
 FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.  
 FT CHAIN 24 112 PROTEIN F2.  
 FT CHAIN 113 550 PROTEIN F1.  
 FT TRANSMEM 113 136 POTENTIAL.  
 FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 495 515 POTENTIAL.  
 FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
 FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 550 AA; 59532 MW; 7AA4FICA82169093 CRC64;  
 Query Match 45.0%; Score 72; DB 1; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.0082;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 LSEIKGVVHRLEGV 17  
 DB 288 LSEIKGVVHRLEGV 302  
 RESULT 5  
 VGLF\_RINDR  
 ID VGLF\_RINDR STANDARD; PRT; 546 AA.  
 AC P41356;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
 DE Fusion glycoprotein F1].  
 GN F.  
 OS Rinderpest virus (strain RBOK) (RDV).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses;  
 OC NCBI\_TaxID=36409;  
 RN [1]

```

Query Match          45.0%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LSEIKGVIVHRLEGV 17
    |||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 2
VGLF_RINDB          STANDARD;          PRT;          546 AA.
ID  VGLF_RINDB
AC  P41360;
DT  01-FEB-1995 (Rel. 31, Last sequence update)
DT  01-FEB-1995 (Rel. 31, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE  Fusion glycoprotein F1].
GN  F.
OS  Rinderpest virus (strain RBT1) (RDV).
OC  Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC  Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX  NCBI_TaxID=39007;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=95088609; PubMed=7996154;
RA  Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT  "Nucleotide sequence comparisons of the fusion protein gene from
RT  virulent and attenuated strains of rinderpest virus.";
RL  J. Gen. Virol. 75:3611-3617(1994).
CC  -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC  MEMBRANES.
CC  -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC  LINKED BY A DISULFIDE BOND.
CC  -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC  FAMILY.
CC  -----
CC  THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
EMBL; Z31656; CAA83482.1; -
DR  PIR; S47300; S47300.
DR  HSP; P04849; LSVF.
DR  InterPro: IPR000776; Fusion_gly.
DR  Pfam; PF00523; fusion_gly; 1.
KW  Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT  SIGNAL 1 19
FT  CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT  CHAIN 20 108 F2 PROTEIN.
FT  CHAIN 109 546 F1 PROTEIN.
FT  DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT  TRANSMEM 109 133 POTENTIAL.
FT  TRANSMEM 484 513 POTENTIAL.
FT  DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT  DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT  CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 518 518 O-LINKED (POTENTIAL).
SQ  SEQUENCE 546 AA; 58418 MW; 38B539B89344F401 CRC64;

Query Match          45.0%; Score 72; DB 1; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.0081;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LSEIKGVIVHRLEGV 17
    |||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 4
VGLF_MEASE          STANDARD;          PRT;          550 AA.
ID  VGLF_MEASE
AC  P08300;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE  Fusion glycoprotein F1].
GN  F.
OS  Measles virus (strain AIK-C) (Subacute sclerose panencephalitis
OS  virus).
OC  Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC  Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX  NCBI_TaxID=36408;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93227570; PubMed=8470368;
RA  Mori T., Sasaki K., Hashimoto H., Makino S.;
RT  "Molecular cloning and complete nucleotide sequence of genomic RNA of
RT  the AIK-C strain of attenuated measles virus.";
RL  Virus Genes 7:67-81(1993).
CC  -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC  MEMBRANES.
CC  -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC  LINKED BY A DISULFIDE BOND.
CC  -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC  FAMILY.
CC  -----
CC  THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
EMBL; S58435; AAB26145.1; -
DR  PIR; E48556; E48556.
DR  HSP; P04849; LSVF.
DR  InterPro: IPR000776; Fusion_gly.
DR  Pfam; PF00523; fusion_gly; 1.
KW  Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT  SIGNAL 1 23
FT  CHAIN 24 550 FUSION GLYCOPROTEIN F0.
FT  CHAIN 24 112 PROTEIN F2.
FT  CHAIN 113 550 PROTEIN F1.
FT  DOMAIN 113 136 POTENTIAL.
FT  DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 495 515 POTENTIAL.
FT  DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
FT  DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT  CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE 550 AA; 59540 MW; AAC4DAB92DE0D938 CRC64;

Query Match          45.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.0082;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LSEIKGVIVHRLEGV 17
    |||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 4
VGLF_MEASE          STANDARD;          PRT;          550 AA.
ID  VGLF_MEASE
AC  P08300;

```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:00:31; Search time 5.75346 Seconds  
(without alignments)  
253.383 Million cell updates/sec

Title: US-09-848-834A-9  
Perfect score: 160  
Sequence: 1 KLLSEIKGVIVHRLGVEGSLHWSYGLRPX 31

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	72	45.0	534	1	VGLF_MEASY P26032 measles vir
2	72	45.0	546	1	VGLF_RINDB P41360 rinderpest
3	72	45.0	550	1	VGLF_MEASA P35973 measles vir
4	72	45.0	550	1	VGLF_MEASE P08300 measles vir
5	71	44.4	546	1	VGLF_RINDR P41356 rinderpest
6	66	41.2	546	1	VGLF_RINDL P10864 rinderpest
7	65	40.6	546	1	VGLF_RINDK P12574 rinderpest
8	65	40.6	662	1	VGLF_CDVO P12569 canine dist
9	64	40.0	529	1	VGLF_MEASI P26031 measles vir
10	64	40.0	631	1	VGLF_PHODV P28886 phocine dis
11	57.5	35.9	91	1	GONL_PIG P4921 sus scrofa
12	57.5	35.9	92	1	GONL_HUMAN P01148 homo sapien
13	57	35.6	90	1	GONL_MOUSE P13562 mus musculus
14	57	35.6	92	1	GONL_RAT P07490 rattus norv
15	54	33.8	67	1	GONL_MACMU P55247 macaca mula
16	53	33.1	508	1	VGLG_IHNV P07923 infectious
17	52	32.5	61	1	GONL_SHEEP Q28588 ovis aries
18	52	32.5	63	1	GONL_MESAU O09163 mesocricetu
19	52	32.5	89	1	GONL_XENLA P45656 xenopus lae
20	52	32.5	92	1	GONL_TUPGB Q95335 tupia glis
21	52	32.5	379	1	PURK_BACSU P12045 bacillus su
22	50.5	31.6	92	1	GONL_CAVPO Q54713 cavia porce
23	50.5	31.6	99	1	GONL_DICLA Q91a0 dicentrarch
24	49	30.6	169	1	CX41_THUOB Q91a0 thunnus obe
25	49	30.6	582	1	SYD_CHLTR Q84546 chlamydia t
26	48.5	30.3	95	1	GONL_MORSA O73812 morone saxa
27	48	30.0	10	1	GONL_ALLMI P37041 alligator m
28	48	30.0	92	1	GONL_CHICK P37042 gallus gall
29	48	30.0	94	1	GONL_HAPBU P51918 haplochromi
30	48	30.0	95	1	GONL_PAGMA P70074 pgarus mayo
31	48	30.0	95	1	GONL_SPAAU P51919 sparus aura
32	48	30.0	124	1	UCN1_HUMAN P55089 homo sapien
33	48	30.0	576	1	DFAL_ANASP Q8ynw5 anabaena sp

34	47	29.4	122	1	UCN1_MOUSE	P81615 mus musculu
35	47	29.4	452	1	HEMF_RHOSH	P33770 rhodobacter
36	47	29.4	452	1	HEMN_RHOSH	P95651 rhodobacter
37	47	29.4	573	1	DFAL_SYNY3	Q55393 synechocyst
38	47	29.4	579	1	DFAL_ANASP	Q82000 anabaena sp
39	46.5	29.1	207	1	TER2_ECOLI	Q04483 escherichia
40	46.5	29.1	417	1	RFL_PYRAB	Q9v151 pyrococcus
41	46	28.7	110	1	YHBJ_ACTAC	P96769 actinobacil
42	46	28.7	555	1	VGLF_P1LHC	P12605 human parai
43	46	28.7	652	1	GSHL_ONCVO	Q9nfn6 onchocerca
44	45.5	28.4	338	1	DCUP_AQUAE	O66667 aquifex ae
45	45.5	28.4	417	1	RFL_PYRHO	O59264 pyrococcus

ALIGNMENTS

RESULT 1  
VGLF\_MEASY  
ID VGLF\_MEASY STANDARD; PRT; 534 AA.  
AC P26032;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
DE Fusion glycoprotein F1].  
GN F.  
OS Measles virus (strain Yamagata-1) (Subacute sclerose panencephalitis  
OS virus).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90385702; PubMed=1698327;  
RA Komase K., Haga T., Yoshikawa Y., Sato T.A., Yamanouchi K.;  
RT "Molecular analysis of structural protein genes of the Yamagata-1  
RT strain of defective subacute sclerosing panencephalitis virus. IV.  
RT Nucleotide sequence of the fusion gene."  
RL Virus Genes 4:173-181(1990).  
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
CC MEMBRANES.  
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
CC LINKED BY A DISULFIDE BOND  
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: D10548; BAA01405.1; -  
CC HSP: P04849; ISVF.  
CC InterPro: IPR000776; Fusion\_gly.  
CC Pfam: PF00523; fusion\_gly; 1.  
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
FT SIGNAL 1 23  
FT CHAIN 24 534 FUSION GLYCOPROTEIN F0.  
FT CHAIN 24 112 PROTEIN F2.  
FT CHAIN 113 534 PROTEIN F1.  
FT TRANSMEM 113 136 POTENTIAL.  
FT DOMAIN 113 494 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 495 515 POTENTIAL.  
FT DOMAIN 516 534 CYTOPLASMIC (POTENTIAL).  
FT DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 534 AA; 57963 MW; F5B21757E643844D CRC64;



C:Superfamily: parainfluenza virus cell fusion protein  
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>  
 F:89-106/Domain: transmembrane #status predicted <TM1>  
 F:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>  
 F:194-219/Domain: transmembrane #status predicted <TM2>  
 F:575-595/Domain: transmembrane #status predicted <TM3>  
 F:110,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.0%; Score 64; DB 1; Length 631;  
 Best Local Similarity 80.0%; Pred. NO. 0.52;  
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGV 17  
 |||:|||||  
 Db 369 LSEIKGVIVHRLGV 383

Search completed: September 24, 2003, 17:49:29  
 Job time : 11.9058 secs

C;Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 24-Nov-1999  
 C;Accession: JQ2223  
 C;Visser, I.K.G.; van der Heijden, R.W.J.; van de Bilt, M.W.G.; Kenter, M.J.H.; Oervell  
 J. Gen. Virol. 74, 1989-1994, 1993  
 A;Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites and  
 e virus entity.  
 A;Reference number: JQ2223; MUID:93389459; PMID:8376973  
 A;Accession: JQ2223  
 A;Molecule type: mRNA  
 A;Residues: 1-542 <VIS>  
 A;Cross-references: GB:L07075  
 A;Note: the authors translated the codon ATC for residue 4 as Leu  
 C;Comment: This fusion protein F0 is cleaved into F1 and F2 chains.  
 C;Genetics:  
 A;Gene: F  
 C;Superfamily: parainfluenza virus cell fusion protein  
 C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F;1-15/Domain: signal sequence #status predicted <SIG>  
 F;16-542/Product: fusion protein #status predicted <MAT>  
 F;16-99/Product: F2 chain #status predicted <F2C>  
 F;105-542/Product: F1 chain #status predicted <F1C>  
 F;105-135/Region: hydrophobic  
 F;486-512/Domain: transmembrane #status predicted <TM>  
 F;21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 40.6%; Score 65; DB 2; Length 542;  
 Best Local Similarity 86.7%; Pred. No. 0.39;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 LSEIKGVIVHRLGV 17  
 III:|||||  
 Db 280 LSEKGVIVHRLGV 294  
 RESULT 12  
 VGNZCD  
 cell fusion glycoprotein precursor - canine distemper virus  
 N;Contains: fusion protein F1; fusion protein F2  
 C;Species: canine distemper virus  
 C;Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Jul-1999  
 C;Accession: J50321  
 R;Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.  
 Virus Res. 8, 373-386, 1987  
 A;Title: The nucleotide sequence of the gene encoding the F protein of canine distemper  
 A;Reference number: J50321; MUID:88129050; PMID:3433924  
 A;Accession: J50321  
 A;Molecule type: mRNA  
 A;Residues: 1-662 <BAR>  
 A;Cross-references: GB:M21849; NID:9323241; PIDN:AAA42878.1; PID:9323242  
 C;Genetics:  
 A;Gene: F  
 C;Superfamily: parainfluenza virus cell fusion protein  
 C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F;1-135/Domain: signal sequence #status predicted <SIG>  
 F;136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>  
 F;225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>  
 F;606-629/Domain: transmembrane #status predicted <MEM>  
 F;62,141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 40.6%; Score 65; DB 1; Length 662;  
 Best Local Similarity 86.7%; Pred. No. 0.39;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 LSEIKGVIVHRLGV 17  
 III:|||||  
 Db 400 LSEKGVIVHRLGV 414  
 RESULT 13  
 S21382  
 cell fusion protein - canine distemper virus  
 C;Species: canine distemper virus  
 C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Nov-1999

C;Accession: S21382  
 R;Wild, T.F.; Bernard, A.; Spehner, D.; Villeval, D.; Drillien, R.  
 submitted to the EMBL Data Library, April 1992  
 A;Description: Vaccination of mice against canine distemper virus induced encephaliti  
 A;Reference number: S21382  
 A;Accession: S21382  
 A;Status: preliminary  
 A;Molecule type: genomic RNA  
 A;Residues: 1-662 <WIL>  
 A;Cross-references: EMBL:X65509; NID:958853; PIDN:CAA46481.1; PID:958854  
 C;Superfamily: parainfluenza virus cell fusion protein  
 Query Match 40.6%; Score 65; DB 2; Length 662;  
 Best Local Similarity 86.7%; Pred. No. 0.39;  
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 LSEIKGVIVHRLGV 17  
 III:|||||  
 Db 400 LSEKGVIVHRLGV 414  
 RESULT 14  
 VGNZPD  
 cell fusion glycoprotein precursor - phocine distemper virus  
 N;Contains: fusion protein F1; fusion protein F2  
 C;Species: phocine distemper virus  
 C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 25-Oct-1996  
 C;Accession: JQ1368  
 R;Koovamees, J.; Blixenkron-Woeller, M.; Sharma, B.; Oervell, C.; Norrby, E.  
 J. Gen. Virol. 72, 2959-2966, 1991  
 A;Title: The nucleotide sequence and deduced amino acid composition of the haemagglut  
 A;Reference number: JQ1368; MUID:92113538; PMID:1765768  
 A;Accession: JQ1368  
 A;Molecule type: genomic RNA  
 A;Residues: 1-631 <NOV>  
 C;Genetics:  
 A;Gene: F  
 C;Superfamily: parainfluenza virus cell fusion protein  
 C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F;1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>  
 F;89-106/Domain: transmembrane #status predicted <TM1>  
 F;189-193/Region: cleavage processing #status predicted  
 F;194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>  
 F;194-212/Domain: transmembrane #status predicted <TM2>  
 F;575-595/Domain: transmembrane #status predicted <TM3>  
 F;110,142,148,486/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 40.0%; Score 64; DB 1; Length 631;  
 Best Local Similarity 80.0%; Pred. No. 0.52;  
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 LSEIKGVIVHRLGV 17  
 III:|||||  
 Db 369 LSEKGVIVHRLGV 383  
 RESULT 15  
 A48346  
 cell fusion glycoprotein precursor - phocine distemper virus (strain Ulster/88)  
 N;Contains: fusion protein F1; fusion protein F2  
 C;Species: phocine distemper virus  
 C;Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 25-Oct-1996  
 C;Accession: A48346  
 R;Curran, M.D.; Lu, Y.J.; Rima, B.K.  
 Arch. Virol. 126, 159-169, 1992  
 A;Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced a  
 A;Reference number: A48346; MUID:92398437; PMID:1524494  
 A;Accession: A48346  
 A;Molecule type: mRNA  
 A;Residues: 1-631 <CUR>  
 A;Note: sequence extracted from NCBI backbone (NCBIN:113098, NCBIP:113099)  
 C;Genetics:  
 A;Gene: F

C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:26-110/Product: cell fusion glycoprotein F2 #status predicted <FF2>  
 F:111-553/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
 F:501-517/Domain: transmembrane #status predicted <TMN>  
 F:32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.0%; Score 72; DB 1; Length 553;  
 Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||:|||||  
 Db 291 LSEIKGVIVHRLEGV 305

RESULT 7  
 VGNZRL  
 cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)  
 N:Contains: fusion glycoprotein F1; fusion glycoprotein F2  
 C:Species: rinderpest virus  
 C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 25-Oct-1996  
 C:Accession: A31051  
 R:Hsu, D.; Yamanaka, M.; Miller, J.; Dale, B.; Grubman, M.; Yilma, T.  
 Virology 166, 149-153, 1988  
 A:Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis  
 A:Reference number: A31051; MUID:88322864; PMID:3413983  
 A:Accession: A31051  
 A:Molecule type: genomic RNA  
 A:Residues: 1-546 <HSU>  
 C:Genetics:  
 A:Gene: F  
 C:Superfamily: parainfluenza virus cell fusion protein  
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-108/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
 F:109-546/Product: cell fusion glycoprotein F2 #status predicted <FF2>  
 F:109-134/Domain: transmembrane #status predicted <TM1>  
 F:491-513/Domain: transmembrane #status predicted <TM2>  
 F:25,57,63,518/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 44.4%; Score 71; DB 1; Length 546;  
 Best Local Similarity 93.3%; Pred. No. 0.046;  
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||:|||||  
 Db 284 LSEIKGVIVHRLEGV 298

RESULT 8  
 S47305  
 gene F protein - rinderpest virus  
 C:Species: rinderpest virus  
 C:Date: 20-Oct-1994 #sequence\_revision 08-Sep-1995 #text\_change 20-Sep-1999  
 C:Accession: S47305; S47301  
 R:Barton, M.D.; Barrett, T.  
 submitted to the EMBL Data Library, March 1994  
 A:Description: The sequence of the N and L genes of Rinderpest virus, and the 50 and 30  
 A:Reference number: S47283  
 A:Accession: S47305  
 A:Molecule type: mRNA  
 A:Residues: 1-546 <BAR>  
 A:Cross-references: EMBL:230697; NID:g535396; PIDN:CAA93181.1; PID:g535401; EMBL:230700;  
 C:Superfamily: parainfluenza virus cell fusion protein  
 C:Keywords: transmembrane protein

Query Match 44.4%; Score 71; DB 2; Length 546;  
 Best Local Similarity 93.3%; Pred. No. 0.046;  
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||:|||||

Db 284 LSEIKGVIVHRLEGV 298

RESULT 9  
 S47034  
 cell fusion protein precursor - porpoise morbillivirus  
 N:Alternate names: F protein  
 C:Species: porpoise morbillivirus  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Nov-1999  
 C:Accession: S47034  
 R:Bolt, G.; Gottschalk, E.; Blixenkrone-Moeller, M.; Wishaupt, R.G.A.; Welsh, M.-.;  
 submitted to the EMBL Data Library, July 1994  
 A:Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbilli  
 A:Reference number: S47034  
 A:Accession: S47034  
 A:Molecule type: mRNA  
 A:Residues: 1-352 <BOL>  
 A:Cross-references: EMBL:X80757; NID:g520639; PIDN:CAA56731.1; PID:g520640  
 A:Experimental source: isolate Ulster 88  
 A:Note: The source is designated as Cetacean morbillivirus  
 C:Superfamily: parainfluenza virus cell fusion protein  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:26-552/Product: fusion protein #status predicted <MAT>

Query Match 41.6%; Score 66.5; DB 2; Length 552;  
 Best Local Similarity 61.5%; Pred. No. 0.2;  
 Matches 16; Conservative 1; Mismatches 2; Indels 7; Gaps 1;

QY 3 LSEIKGVIVHRLEGVGFSLHWSYGL 28  
 |||||:|||||  
 Db 290 LSEIKGVIVHRLEAV-----SYNL 308

RESULT 10  
 VGNZRL  
 cell fusion glycoprotein precursor - rinderpest virus (strain L)  
 N:Contains: fusion glycoprotein F1; fusion glycoprotein F2  
 C:Species: rinderpest virus  
 C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 16-Jul-1999  
 C:Accession: A28921  
 R:Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.  
 Virology 164, 523-530, 1988  
 A:Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence  
 A:Reference number: A28921; MUID:88219541; PMID:3285575  
 A:Accession: A28921  
 A:Molecule type: mRNA  
 A:Residues: 1-546 <TSU>  
 A:Cross-references: GB:M20870; NID:g333898; PIDN:AAA47399.1; PID:g333899  
 C:Genetics:  
 A:Gene: F  
 C:Superfamily: parainfluenza virus cell fusion protein  
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>  
 F:105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>  
 F:109-133/Domain: transmembrane #status predicted <TM1>  
 F:485-513/Domain: transmembrane #status predicted <TM2>  
 F:25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.2%; Score 66; DB 1; Length 546;  
 Best Local Similarity 93.3%; Pred. No. 0.23;  
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||:|||||  
 Db 284 LSEIKGVIVHRLESV 298

RESULT 11  
 JO2223  
 cell fusion protein F0 precursor - phocine distemper virus  
 N:Contains: F1 and F2 chains  
 C:Species: phocine distemper virus

## RESULT 3

JU0274  
cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain Y  
N:Contains: fusion glycoprotein F1; fusion glycoprotein F2  
C:Species: subacute sclerosing panencephalitis virus, SSPV  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jun-2000  
C:Accession: JU0274  
R:Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.  
Virus Genes 4, 173-181, 1990  
A:Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe  
A:Reference number: JU0274; MUID:90385702; PMID:1698327  
A:Accession: JU0274  
A:Molecule type: mRNA  
A:Residues: 1-534 <KOM>  
A:Cross-references: EMBL:D10548; NID:q222256; PIDN:BAA01405.1; PID:q222257  
A:Note: the authors translated the codon GTA for residue 459 as Gly and GGG for residue  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FE2>  
F:108-534/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
F:498-514/Domain: transmembrane #status predicted <TMM>  
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.0%; Score 72; DB 1; Length 534;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17  
|||||  
Db 288 LSEIKGVIVHRLEGV 302

## RESULT 4

S47300  
gene F protein - rinderpest virus  
C:Species: rinderpest virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Oct-1999  
C:Accession: S47300; PQ0865  
R:Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.  
submitted to the EMBL Data Library, March 1994  
A:Description: The complete nucleotide sequence of the fusion protein gene of the vaccin  
A:Reference number: S47299  
A:Accession: S47300  
A:Molecule type: DNA  
A:Residues: 1-546 <EVA>  
A:Cross-references: EMBL:Z31656; NID:g535406; PIDN:CAA83482.1; PID:g535407  
R:Chamberlain, R.W.; Wanwayi, H.M.; Hockley, E.; Shailla, M.S.; Goatley, L.; Knowles, N.J.  
J. Gen. Virol. 74, 2775-2780, 1993  
A:Title: Evidence for different lineages of rinderpest virus reflecting their geographic  
A:Reference number: PQ0865; MUID:94103786; PMID:8277286  
A:Accession: PQ0865  
A:Molecule type: mRNA  
A:Residues: 86-191 <CHA>  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein

Query Match 45.0%; Score 72; DB 2; Length 546;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17  
|||||  
Db 284 LSEIKGVIVHRLEGV 298

## RESULT 5

C:Superfamily: parainfluenza virus cell fusion protein

## E48556

cell fusion glycoprotein precursor - measles virus (strain AIK-C)  
C:Species: measles virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 16-Jul-1999  
C:Accession: E48556  
R:Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.  
Virus Genes 7, 67-81, 1993  
A:Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the AIK  
A:Reference number: A48556; MUID:93227570; PMID:8470368  
A:Accession: E48556  
A:Molecule type: genomic RNA  
A:Residues: 1-550 <MOR>  
A:Cross-references: GB:S58435; NID:g299460; PIDN:AAB26145.1; PID:g299465  
A:Note: sequence extracted from NCBI Backbone (NCBIN:129264; NCBI:P:129272)  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>  
F:108-550/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
F:113-138/Region: hydrophobic  
F:495-514/Domain: transmembrane #status predicted <TMM>  
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.0%; Score 72; DB 1; Length 550;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGV 17  
|||||  
Db 288 LSEIKGVIVHRLEGV 302

## RESULT 6

VGNZMV  
cell fusion glycoprotein precursor - measles virus  
C:Species: measles virus  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1989 #text\_change 16-Jun-2000  
C:Accession: A26962; A25616; PQ0380; PQ0384  
R:Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.  
J. Gen. Virol. 68, 1695-1703, 1987  
A:Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and co  
A:Reference number: A92794; MUID:87224816; PMID:5585281  
A:Accession: A26962  
A:Molecule type: mRNA  
A:Residues: 1-553 <BUC>  
A:Cross-references: GB:D00090; NID:q222061; PIDN:BAA00056.1; PID:q222062  
A:Experimental source: strain Halle  
R:Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Englund, G.; Bellini  
Virology 155, 508-523, 1986  
A:Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles v  
A:Reference number: A94350; MUID:87071668; PMID:3788062  
A:Accession: A25616  
A:Molecule type: mRNA  
A:Residues: 4-553 <RIC>  
A:Cross-references: GB:M14915; NID:g331762; PIDN:AAA46423.1; PID:g331763  
A:Experimental source: strain Edmonston  
R:Schulz, T.F.; Hoar, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.  
J. Gen. Virol. 73, 1581-1586, 1992  
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compar  
A:Reference number: PQ0374; MUID:92300360; PMID:1607874  
A:Accession: PQ0380  
A:Molecule type: genomic RNA  
A:Residues: 272-553 <SCH1>  
A:Experimental source: isolate CL  
A:Accession: PQ0384  
A:Molecule type: genomic RNA  
A:Residues: 272-553 <SCH2>  
A:Experimental source: isolate SE  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:30:06 : Search time 10.9058 Seconds  
(without alignments)  
273.361 Million cell updates/sec

Title: US-09-848-834A-9  
Perfect score: 160  
Sequence: 1 KLLSEIKGVIVHRLEGVEGSLHWSYGLRPX 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	72	45.0	282	PQ0376	cell fusion glycoprotein
2	72	45.0	282	PQ0388	cell fusion glycoprotein
3	72	45.0	534	JU0274	cell fusion glycoprotein
4	72	45.0	546	S47300	gene F protein - r
5	72	45.0	550	E48556	cell fusion glycoprotein
6	72	45.0	553	VGNZMV	cell fusion glycoprotein
7	71	44.4	546	VGNZRK	cell fusion glycoprotein
8	71	44.4	546	S47305	gene F protein - r
9	66.5	41.6	552	S47034	cell fusion glycoprotein
10	66	41.2	546	VGNZRL	cell fusion glycoprotein
11	65	40.6	542	JO2223	cell fusion glycoprotein
12	65	40.6	662	VGNZCD	cell fusion glycoprotein
13	65	40.6	662	S21382	cell fusion glycoprotein
14	64	40.0	631	VGNZPD	cell fusion glycoprotein
15	64	40.0	631	A48346	cell fusion glycoprotein
16	60	37.5	546	S55386	cell fusion glycoprotein
17	57.5	35.9	92	RHHUG	gonadoliberin prec
18	57	35.6	90	RHMSG	gonadoliberin prec
19	57	35.6	92	RHRTG	gonadoliberin prec
20	56	35.0	636	S47299	gene F protein - r
21	54	33.8	67	I78541	gonadoliberin prec
22	53	33.1	508	VGVNFR	spike glycoprotein
23	52	32.5	10	RHPGG	gonadoliberin - pi
24	52	32.5	10	RHSHG	gonadoliberin - sh
25	52	32.5	89	I51423	gonadoliberin prec
26	52	32.5	379	DCBSPK	phosphoribosylamin
27	50.5	31.6	190	T37168	probable tetra-fam
28	50	31.2	333	T23151	hypothetical prote
29	49	30.6	451	AH0063	conserved hypothet

30	49	30.6	582	2	G71500	aspartate-TRNA lig
31	48.5	30.3	393	2	D83589	glutaryl-CoA dehyd
32	48.5	30.3	413	2	B70907	hypothetical prote
33	48	30.0	10	1	RHAQ1	gonadoliberin I -
34	48	30.0	92	2	I50644	gonadoliberin I pr
35	48	30.0	98	2	I50739	gonadoliberin-relea
36	48	30.0	502	2	T36589	probable transmem
37	48	30.0	576	2	AF2361	flavoprotein (impo
38	48	30.0	856	2	I58411	protein-tyrosine k
39	47.5	29.7	222	1	E71024	probable transport
40	47.5	29.7	229	2	JC7219	nuclear protein SR
41	47.5	29.7	586	2	T08293	hypothetical prote
42	47	29.4	339	2	B97755	hypothetical prote
43	47	29.4	388	2	C72710	probable fmu prote
44	47	29.4	452	2	T10882	coproporphyrinogen
45	47	29.4	516	2	T00514	cytochrome P450 ho

## ALIGNMENTS

RESULT 1  
PQ0376  
cell fusion glycoprotein - measles virus (strain TT) (fragment)  
C:Species: measles virus  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 24-Nov-1999  
C:Accession: PQ0376  
R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.  
J. Gen. Virol. 73, 1581-1586, 1992  
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compa-  
A:Reference number: PQ0374; MUID:92300360; PMID:1607874  
A:Accession: PQ0376  
A:Molecule type: genomic RNA  
A:Residues: 1-282 <SCH>  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion

Query Match 45.0%; Score 72; DB 2; Length 282;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
DB 20 LSEIKGVIVHRLEGV 34

RESULT 2  
PQ0388  
cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)  
C:Species: measles virus  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 24-Nov-1999  
C:Accession: PQ0388  
R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.  
J. Gen. Virol. 73, 1581-1586, 1992  
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compa-  
A:Reference number: PQ0374; MUID:92300360; PMID:1607874  
A:Accession: PQ0388  
A:Molecule type: genomic RNA  
A:Residues: 1-282 <SCH>  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion

Query Match 45.0%; Score 72; DB 2; Length 282;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
DB 20 LSEIKGVIVHRLEGV 34



```

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-9301V;
RX MEDLINE-98440529; PubMed=9765410;
RA Takeda M., Kato A., Kobune F., Sakata H., Li Y., Shioda T., Sakai Y.,
RA Asakawa M., Nagai Y.;
RT "Measles virus attenuation associated with transcriptional impediment
RT and a few amino acid changes in the polymerase and accessory
RT proteins."
RL J. Virol. 72:8690-8696(1998).
DR EMBL; AB012949; BAA33877.1; -.
DR EMBL; AB012948; BAA33871.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59512 MW; 7AA4F1D117197BF9 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
Db 288 LSEIKGVIVHRLEGV 302

RESULT 14
Q9QEW8
ID Q9QEX1 PRELIMINARY; PRT; 550 AA.
AC Q9QEX1;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Masusako;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179430; AAF02695.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59559 MW; 609EE024A7E59C54 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
Db 288 LSEIKGVIVHRLEGV 302

RESULT 15
Q9QEW8
ID Q9QEW8 PRELIMINARY; PRT; 550 AA.

```

```

AC Q9QEW8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179437; AAF02702.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59315 MW; 086E51FED235EBBA CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
Db 288 LSEIKGVIVHRLEGV 302

Search completed: September 24, 2003, 17:47:15
Job time : 32.0554 secs

```

AC Q9QEW7; 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OSA-2;  
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,  
 RA Ogura H.;  
 RT "Nucleotide sequences of the fusion protein gene of subacute  
 RT sclerosing panencephalitis viruses: deduced amino acid sequences  
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or  
 RT predicted secondary structure changed.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF179438; AAF02703.1; -;  
 DR HSP; P04849; ISVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly. 1.  
 DR SEQUENCE 550 AA; 59333 MW; 086E51FED5582BBA CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.065;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||  
 DB 288 LSEIKGVIVHRLEGV 302

## RESULT 10

Q9WMK4  
 ID Q9WMK4 PRELIMINARY; PRT; 550 AA.  
 AC Q9WMK4; 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WTF;  
 RX MEDLINE=99329215; PubMed=10400788;  
 RA Johnston I.C., Ter Meulen V., Schneider-Schaulies J.,  
 RA Schneider-Schaulies S.;  
 RT "A recombinant measles vaccine virus expressing wild-type  
 RT glycoproteins : consequences for viral spread and cell tropism.";  
 RL J. Virol. 73:6903-6915(1999).  
 DR EMBL; AJ133108; CAB38075.1; -;  
 DR HSP; P04849; ISVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly. 1.  
 DR SEQUENCE 550 AA; 59580 MW; 82554999685D862 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.065;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||  
 DB 288 LSEIKGVIVHRLEGV 302

## RESULT 11

Q89495  
 ID Q89495 PRELIMINARY; PRT; 550 AA.  
 AC Q89495; 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE Fusion protein.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=92230209; PubMed=1566568;  
 RA Rota J.S., Hummel K.B., Rota P.A., Bellini W.J.;  
 RT "Genetic variability of the glycoprotein genes of current wild-type  
 RT measles isolates.";  
 RL Virology 188:135-142(1992).  
 DR EMBL; M81903; AAA46422.1; -;  
 DR EMBL; M81901; AAA46421.1; -;  
 DR HSP; P04849; ISVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly. 1.  
 DR SEQUENCE 550 AA; 59564 MW; A78EEC9CD6268E58 CRC64;

ID Q89495 PRELIMINARY; PRT; 550 AA.  
 AC Q89495; 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=92230209; PubMed=1566568;  
 RA Rota J.S., Hummel K.B., Rota P.A., Bellini W.J.;  
 RT "Genetic variability of the glycoprotein genes of current wild-type  
 RT measles isolates.";  
 RL Virology 188:135-142(1992).  
 DR EMBL; M81903; AAA46422.1; -;  
 DR EMBL; M81901; AAA46421.1; -;  
 DR HSP; P04849; ISVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly. 1.  
 DR SEQUENCE 550 AA; 59564 MW; A78EEC9CD6268E58 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.065;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||  
 DB 288 LSEIKGVIVHRLEGV 302

## RESULT 12

Q8V049  
 ID Q8V049 PRELIMINARY; PRT; 550 AA.  
 AC Q8V049; 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=G954;  
 RX MEDLINE=21635526; PubMed=11773423;  
 RA Waku Koumou D., Wild T.F.;  
 RT "Adaptation of wild-type measles virus to tissue culture.";  
 RL J. Virol. 76:1505-1509(2002).  
 DR EMBL; AY059392; AAL29688.1; -;  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly. 1.  
 DR SEQUENCE 550 AA; 59551 MW; 9A7A4BA99E4DA8E9 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.065;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||  
 DB 288 LSEIKGVIVHRLEGV 302

## RESULT 13

Q9YJ94  
 ID Q9YJ94 PRELIMINARY; PRT; 550 AA.  
 AC Q9YJ94; 01-MAY-1999 (TrEMBLrel. 10, Created)

```

RT fusogenicity of measles virus."
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGAHATA;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; D63926; BAA09958.1; -.
DR EMBL; AF179431; AAF02696.1; -.
DR HSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59530 MW; 97C991C7E2169839 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
Db 288 LSEIKGVIVHRLEGV 302

RESULT 6
Q9QEX0
ID Q9QEX0 PRELIMINARY; PRT; 550 AA.
AC Q9QEX0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OX NCBI_TaxID=11234;
RN [1]
RC STRAIN=Toyoshima;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179432; AAF02697.1; -.
DR HSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59504 MW; 2AA969D37FA5CA17 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
Db 288 LSEIKGVIVHRLEGV 302

RESULT 7
Q9QEW9
ID Q9QEW9 PRELIMINARY; PRT; 550 AA.
AC Q9QEW9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Fusion protein.

```

```

OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OX NCBI_TaxID=11234;
RN [1]
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179436; AAF02701.1; -.
DR HSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59405 MW; 0AB6DBFC5DD22BBA CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
Db 288 LSEIKGVIVHRLEGV 302

RESULT 8
P90330
ID P90330 PRELIMINARY; PRT; 550 AA.
AC P90330;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OX NCBI_TaxID=11234;
RN [1]
RC STRAIN=Nagahata(HB);
RA Sheng J., Watanabe M., Ueda S.;
RT "Selection of a neurotropic variant of measles virus.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagahata(HB);
RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT "An amino acid alteration of F protein responsible for the enhanced
RT fusogenicity of measles virus.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; D63924; BAA09951.1; -.
DR HSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59589 MW; 73E7BD457ABA39B7 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17
Db 288 LSEIKGVIVHRLEGV 302

RESULT 9
Q9QEW7
ID Q9QEW7 PRELIMINARY; PRT; 550 AA.

```

## RESULT 2

Q04242 ID Q04242 PRELIMINARY; PRT; 537 AA.  
 AC Q04242;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89003063; PubMed=3167982;  
 RA Cattaneo R., Schmid A., Eschle D., Baczko K., ter Meulen V.,  
 RA Billeter M.A.;  
 RT "Biased hypermutation and other genetic changes in defective measles  
 RT viruses in human brain infections.";  
 RL Cell 55:255-265(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Cattaneo R., Billeter M.A.;  
 RL Virology 0:0-0(0).  
 DR EMBL; X16567; CAA34574.1; -;  
 DR EMBL; X16567; CAA34575.1; -;  
 DR HSSP; P04849; ISVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 537 AA; 58275 MW; D0A60AC6D979E06 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 537;  
 Best Local Similarity 100.0%; Pred. No. 0.063;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||  
 DB 291 LSEIKGVIVHRLEGV 305

## RESULT 3

Q9PXA4 ID Q9PXA4 PRELIMINARY; PRT; 545 AA.  
 AC Q9PXA4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OSA-3;  
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,  
 RA Ogura H.;  
 RT "Nucleotide sequences of the fusion protein gene of subacute  
 RT sclerosing panencephalitis viruses: deduced amino acid sequences  
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or  
 RT predicted secondary structure changed.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF179440; AAF02705.1; -;  
 DR EMBL; AF179439; AAF02704.1; -;  
 DR HSSP; P04849; ISVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 545 AA; 58907 MW; 0234C28AE193E77D CRC64;

Query Match 45.0%; Score 72; DB 12; Length 545;

Best Local Similarity 100.0%; Pred. No. 0.064; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||  
 DB 288 LSEIKGVIVHRLEGV 302

## RESULT 4

Q91HA5 ID Q91HA5 PRELIMINARY; PRT; 546 AA.  
 AC Q91HA5;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Rinderpest virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11241;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K;  
 RX MEDLINE=21014265; PubMed=11186456;  
 RA Aianot P.K., Smeinev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,  
 RA Gusev A.A.;  
 RT "Primary structure of the F-gene from Rinderpest virus strain K.";  
 RL Mol. Gen. Mikrobiol. Virusol. 4:29-33(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K;  
 RA Aianot P.K., Smeinev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,  
 RA Gusev A.A.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF035887; AAK63190.1; -;  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 546 AA; 58572 MW; 449B2B2DD7405F0B CRC64;

Query Match 45.0%; Score 72; DB 12; Length 546;  
 Best Local Similarity 100.0%; Pred. No. 0.064;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGV 17  
 |||||  
 DB 284 LSEIKGVIVHRLEGV 298

## RESULT 5

P90331 ID P90331 PRELIMINARY; PRT; 550 AA.  
 AC P90331;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NAGAHATA;  
 RA Sheng J., Watanabe M., Ueda S.;  
 RT "Selection of a neurotropic variant of measles virus.";  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NAGAHATA;  
 RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;  
 RT "An amino acid alteration of F protein responsible for the enhanced

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:29:11 ; Search time 30.0554 Seconds  
(without alignments)  
286.163 Million cell updates/sec

Title: US-09-848-834a-9  
Perfect score: 160  
Sequence: 1 KLLSEIKGVIVHRLGEGVPSLHWSYGLRPX 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-rvirus:\*  
16: sp-bacteriap:\*  
17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	45.0	534	12	Q04243 measles vir
2	72	45.0	537	12	Q04242 measles vir
3	72	45.0	545	12	Q9PXA4 measles vir
4	72	45.0	546	12	Q91HA5 rinderpest
5	72	45.0	550	12	P90331 measles vir
6	72	45.0	550	12	Q9QEX0 measles vir
7	72	45.0	550	12	Q9QEW9 measles vir
8	72	45.0	550	12	P90330 measles vir
9	72	45.0	550	12	Q9QEW7 measles vir
10	72	45.0	550	12	Q9WNK4 measles vir
11	72	45.0	550	12	Q89495 measles vir
12	72	45.0	550	12	Q8V049 measles vir
13	72	45.0	550	12	Q9YJ94 measles vir
14	72	45.0	550	12	Q9QEX1 measles vir
15	72	45.0	550	12	Q9QEW8 measles vir
16	72	45.0	553	12	Q93055 measles vir

17	72	45.0	553	12	Q9IC36	Q91C36 measles vir
18	72	45.0	553	12	P88973	P88973 measles vir
19	72	45.0	553	12	Q83536	Q83536 measles vir
20	72	45.0	553	12	O11383	O11383 measles vir
21	72	45.0	553	12	Q91FK2	Q91FK2 measles vir
22	72	45.0	553	12	Q83533	Q83533 measles vir
23	72	45.0	553	12	Q83525	Q83525 measles vir
24	72	45.0	553	12	Q83518	Q83518 measles vir
25	72	45.0	553	12	P88974	P88974 measles vir
26	72	45.0	553	12	Q83527	Q83527 measles vir
27	72	45.0	553	12	Q83521	Q83521 measles vir
28	72	45.0	553	12	Q83530	Q83530 measles vir
29	72	45.0	553	12	O91248	O91248 measles vir
30	72	45.0	553	12	Q910P2	Q910P2 measles vir
31	72	45.0	553	12	Q04244	Q04244 measles vir
32	72	45.0	579	12	Q9PW04	Q9PW04 measles vir
33	68	42.5	545	12	Q9QEW6	Q9QEW6 measles vir
34	68	42.5	553	12	O11380	O11380 measles vir
35	66.5	41.6	552	12	Q66147	Q66147 cetacean mo
36	65	40.6	528	12	Q9YJW9	Q9YJW9 canine dist
37	65	40.6	530	12	Q8QV06	Q8QV06 canine dist
38	65	40.6	662	12	Q9DX22	Q9DX22 canine dist
39	65	40.6	662	12	Q91KN3	Q91KN3 canine dist
40	65	40.6	662	12	Q9YK17	Q9YK17 canine dist
41	65	40.6	662	12	O89327	O89327 canine dist
42	63.5	39.7	552	12	Q66409	Q66409 dolphin mor
43	63.5	39.7	552	12	O56852	O56852 dolphin mor
44	61	38.1	553	12	Q83629	Q83629 measles vir
45	60	37.5	546	12	O84926	O84926 peste-des-p

## ALIGNMENTS

## RESULT 1

Q04243 ID Q04243 PRELIMINARY; PRT; 534 AA.  
AC Q04243;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)  
DE Fusion protein.  
GN F.  
OS Measles virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11234;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=69003063; PubMed=3167982;  
RA Cattaneo R., Schmid A., Eschle D., Bacsko K., ter Meulen V.,  
RA Billeter M.A.;  
RT "Biased hypermutation and other genetic changes in defective measles  
RT viruses in human brain infections.";  
RL Cell 55:255-265(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Cattaneo R., Billeter M.A.;  
RL Virology 0:0-0(0).  
DR EMBL; X16568; CAA34581.1; -;  
DR EMBL; X16568; CAA34582.1; -;  
DR HSSP; P04849; 1SVF.  
DR InterPro; IPR000776; Fusion\_gly.  
DR Pfam; PF00523; fusion\_gly; 1.  
SQ SEQUENCE 534 AA; 57899 MW; 637245E23B5BE044 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 534;

Best Local Similarity 100.0%; Pred. No. 0.063;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LSEIKGVIVHRLGCV 17  
|||||

Db 291 LSEIKGVIVHRLGCV 305



```

QY      15 SLFNNFTVSFWLRVVK 30
Db      935 SNYENFTSFWIRPK 950

RESULT 4
BXF_CLOBO STANDARD; PRT; 1274 AA.
AC      P30996;
DT      01-JUL-1993 (Rel. 26, Created)
DT      01-JUL-1993 (Rel. 26, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
DE      (Bontoxilysin F).
GN      BOTF
OS      Clostridium botulinum.
OC      Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC      Clostridium.
OX      NCBI_TaxID=1491;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 23387;
RX      MEDLINE=93012902; PubMed=1398040;
RA      East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA      Roberts T.A., Thompson D.E.;
RT      "Sequence of the gene encoding type F neurotoxin of Clostridium
RT      botulinum";
RT      J. Clin. Microbiol. 31:2255-2262(1993).
RL      FEMS Microbiol. Lett. 75:225-230(1992).
RN      [2]
RP      SEQUENCE OF 1-64 FROM N.A.
RC      STRAIN=Hobbs FT10;
RX      MEDLINE=94297488; PubMed=7764998;
RA      East A.K., Collins M.D.;
RT      "Conserved structure of genes encoding components of botulinum
RT      neurotoxin complex M and the sequence of the gene coding for the
RT      nontoxic component in nonproteolytic Clostridium botulinum type F.";
RT      Curr. Microbiol. 29:69-77(1994).
RN      [3]
RP      SEQUENCE OF 634-1002 FROM N.A.
RX      MEDLINE=94013372; PubMed=8408542;
RA      Campbell K., East A.K., Collins M.D.;
RT      "Gene probes for identification of the botulin neurotoxin gene and
RT      specific identification of neurotoxin types B, E, and F.";
RT      J. Clin. Microbiol. 31:2255-2262(1993).
RN      [4]
RP      IDENTIFICATION OF SUBSTRATE.
RX      MEDLINE=94230352; PubMed=8175689;
RA      Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA      Roques B., Fyke E.M., Suedhof T.C., Jahn R., Niemann H.;
RT      "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT      F botulin neurotoxins and tetanus toxin.";
RT      J. Biol. Chem. 269:12764-12772(1994).
CC      -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC      RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC      AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC      WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC      INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC      ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-1-LYS-59
CC      BOND OF SYNAPTOSOMAL-1 AND -2.
CC      -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC      neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC      detected action on small molecule substrates.
CC      -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC      -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC      HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC      WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC      FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC      BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC      -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M92906; AAA23263.1; -
DR      EMBL; S73676; AAC60475.1; -
DR      EMBL; X70820; CAA50151.1; -
DR      EMBL; X70816; CAA50147.1; -
DR      PIR; I40813; I40813.
DR      PIR; S48109; S48109.
DR      HSSP; P10845; 3BTA.
DR      MEROPS; M27.002; -
DR      InterPro; IPR000395; Bontoxilysin.
DR      InterPro; IPR006025; Zn_Mtpeptidse.
DR      Pfam; PF01742; Peptidase_M27; 1.
DR      ProDom; PD001963; Bontoxilysin; 1.
DR      ProSite; PS00142; ZINC_PROTEASE; 1.
KW      Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT      CHAIN 1 436
FT      METAL 437 1274
FT      ACT_SITE 227 228
FT      ACT_SITE 228 228
FT      METAL 231 231
FT      DISULFID 429 445
FT      SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;
SQ      QUERY MATCH 31.5%; Score 63; DB 1; Length 1274;
SQ      Best Local Similarity 56.2%; Pred. No. 0.95;
SQ      Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY      15 SLFNNFTVSFWLRVVK 30
Db      928 SRYQNFSISFWIRPK 943

RESULT 5
BXB_CLOBO STANDARD; PRT; 1290 AA.
AC      P10844; P10843;
DT      01-JUL-1989 (Rel. 11, Created)
DT      01-JUL-1993 (Rel. 26, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
DE      (Bontoxilysin B).
GN      BOTB
OS      Clostridium botulinum.
OC      Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC      Clostridium.
OX      NCBI_TaxID=1491;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=92384550; PubMed=1514783;
RA      Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA      Minton N.P.;
RT      "Molecular cloning of the Clostridium botulinum structural gene
RT      encoding the type B neurotoxin and determination of its entire
RT      nucleotide sequence.";
RT      Appl. Environ. Microbiol. 58:2345-2354(1992).
RN      [2]
RP      SEQUENCE OF 35-245 FROM N.A.
RC      STRAIN=NC7C 7273;
RA      Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
RL      Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE OF 633-993 FROM N.A.
RC      STRAIN=NC7C 7273;
RX      MEDLINE=94013372; PubMed=8408542;
RA      Campbell K., East A.K., Collins M.D.;
RT      "Gene probes for identification of the botulin neurotoxin gene and
RT      specific identification of neurotoxin types B, E, and F.";

```

RL FERS Lett. 376:41-44(1995).  
RN [6]  
RP SEQUENCE OF 1-16.  
RX MEDLINE=84178501; PubMed=6370252;  
RA Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;  
RT "Partial amino acid sequence of the heavy and light chains of  
FT botulinum neurotoxin type A.";  
RL Biochem. Biophys. Res. Commun. 119:900-904(1984).  
RN [7]  
RP SEQUENCE OF 1-46.  
RA Dasgupta B.R., Foley J., Niece R.;  
RT "Partial sequence of the light chain of botulinum neurotoxin type A.";  
RL Biochemistry 26:4162-4162(1987).  
RN [8]  
RP SEQUENCE OF 1-5 AND 44-456.  
RX MEDLINE=91120847; PubMed=2126206;  
RA Dasgupta B.R., Dekleva M.L.;  
RT "Botulinum neurotoxin type A: sequence of amino acids at the  
FT N-terminus and around the nicking site.";  
RL Biochimie 72:661-664(1990).  
RN [9]  
RP SEQUENCE OF 448-464 AND 872-895.  
RX MEDLINE=89024662; PubMed=3178218;  
RA Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;  
RT "Botulinum neurotoxin type A: cleavage of the heavy chain into two  
FT halves and their partial sequences.";  
RL Arch. Biochem. Biophys. 266:142-151(1988).  
RN [10]  
RP SEQUENCE OF 448-482.  
RX MEDLINE=85285016; PubMed=3896784;  
RA Shone C.C., Hambleton P., Welling J.;  
RT "Inactivation of Clostridium botulinum type A neurotoxin by trypsin  
FT and purification of two tryptic fragments. Proteolytic action near  
FT the COOH-terminus of the heavy subunit destroys toxin-binding  
FT activity.";  
RL Eur. J. Biochem. 151:75-82(1985).  
RN [11]  
RP IDENTIFICATION OF SUBSTRATE.  
RX MEDLINE=94063091; PubMed=8243676;  
RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,  
RA Benfenati F., Wilson M.C., Montecucco C.;  
RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct  
FT COOH-terminal peptide bonds.";  
RL FERS Lett. 335:99-103(1993).  
RN [12]  
RP IDENTIFICATION OF SUBSTRATE.  
RX MEDLINE=94124495; PubMed=8294407;  
RA Binz T., Blas J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,  
RA Jahn R., Niemann H.;  
RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";  
RL J. Biol. Chem. 269:1617-1620(1994).  
RN [13]  
RP MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.  
RX MEDLINE=21556941; PubMed=11700044;  
RA Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.;  
RT "Site-directed mutagenesis identifies active-site residues of the  
FT light chain of botulinum neurotoxin type A.";  
RL Biochem. Biophys. Res. Commun. 288:1231-1237(2001).  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).  
RX MEDLINE=98455071; PubMed=9783750;  
RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;  
RT "Crystal structure of botulinum neurotoxin type A and implications  
FT for toxicity.";  
RL Nat. Struct. Biol. 5:898-902(1998).  
CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin  
CC binds with high affinity to peripheral neuronal presynaptic  
CC membrane, is then internalized by receptor-mediated endocytosis.  
CC The C-terminus of the heavy chain (H) is responsible for the  
CC adherence of the toxin to the cell surface while the N-terminus  
CC mediates transport of the light chain from the endocytic vesicle  
CC to the cytosol. After translocation, the light chain (L)  
CC hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking

neurotransmitter release. Inhibition of acetylcholine release  
results in flaccid paralysis, with frequent heart or respiratory  
failure.  
-1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No  
detected action on small molecule substrates.  
-1- COFACTOR: Binds 1 zinc ion per subunit.  
-1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a  
heavy chain (H).  
-1- SUBCELLULAR LOCATION: Secreted.  
-1- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for  
the treatment of strabismus and blepharospasm associated with  
dystonia and cervical dystonia. Also used for the treatment of  
hemifacial spasm and a number of other neurological disorders  
characterized by abnormal muscle contraction.  
-1- MISCELLANEOUS: There are seven antigenically distinct forms of  
botulinum neurotoxin: Types A, B, C1, D, E, F, and G.  
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
-1- DATABASE: NAME-BOTOX product information Web site:  
WWW="http://www.botox.com/index.jsp?hpsproductinfo".  
-1- DATABASE: NAME-Protein Spotlight;  
NOTE=Issue 19 of February 2002;  
WWW="http://www.expasy.org/spotlight/articles/spl0109.html".  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
EMBL; X52066; CAA36289.1; -  
DR EMBL; M30196; AAA23262.1; -  
DR EMBL; X2973; CAA63551.1; -  
DR EMBL; D67030; BAA11051.1; -  
DR EMBL; M27892; AAA23269.1; -  
DR PIR; A35294; BTCLAB.  
DR PDB; 3BTA; 01-OCT-99.  
DR MEROPS; M27.002; -  
DR InterPro; IPR000395; Bontoxilysin.  
DR InterPro; IPR006025; Zn\_Mtpeptidse.  
DR Pfam; PF01742; Peptidase\_M27; 1.  
DR ProDom; PD001963; Bontoxilysin; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;  
KW Pharmaceutical; 3D-structure.  
FT INIT\_MET 0  
FT CHAIN 1 447: BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.  
FT CHAIN 448 1295: BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.  
FT METAL 222 222 ZINC (CATALYTIC).  
FT ACT\_SITE 223 223 ZINC (CATALYTIC).  
FT METAL 226 226 ZINC (CATALYTIC).  
FT METAL 261 261 ZINC (CATALYTIC).  
FT DISULFID 429 453 INTERCHAIN.  
FT DISULFID 1234 1279 POTENTIAL.  
FT TRANSMEM 626 646 POTENTIAL.  
FT TRANSMEM 655 675 V -> A.  
FT VARIANT 26 26 E->A: DRASTIC DECREASE IN ENZYMIC  
FT MUTAGEN 261 261 ACTIVITY.  
FT MUTAGEN 265 265 F->A: DECREASE IN ENZYMIC ACTIVITY.  
FT MUTAGEN 365 365 Y->A: DECREASE IN ENZYMIC ACTIVITY.  
FT CONFLICT 1 1 P -> Q (IN REF. 1).  
FT CONFLICT 479 479 E -> P (IN REF. 9).  
FT CONFLICT 875 875 T -> L (IN REF. 8).  
FT CONFLICT 891 891 S -> K (IN REF. 8).  
SQ SEQUENCE 1295 AA; 149322 MW; 858342F754862579 CRC64;  
Query Match 32.5%; Score 65; DB 1; Length 1295;  
Best Local Similarity 56.2%; Pred. No. 0.51;  
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;



RL Eur. J. Biochem. 188:39-45(1990).  
 RN [6]  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE=92037649; PubMed=1935979;  
 RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;  
 RT "Limited proteolysis of tetanus toxin. Relation to activity and  
 identification of cleavage sites.";  
 RL Eur. J. Biochem. 202:41-51(1991).  
 RN [7]  
 RP IDENTIFICATION AS ZINC-PROTEASE.  
 RX MEDLINE=93010948; PubMed=1396558;  
 RA Schiavo G., Benfenati F., Rossetto O., Benfenati F., Tauc L.,  
 RA Montecucco C.;  
 RT "Tetanus toxin is a zinc protein and its inhibition of  
 neurotransmitter release and protease activity depend on zinc.";  
 RL EMBO J. 11:3577-3583(1992).  
 RN [8]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=93063293; PubMed=1331807;  
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,  
 RA Dasgupta B.R., Montecucco C.;  
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release  
 by proteolytic cleavage of synaptobrevin.";  
 RL Nature 359:832-835(1992).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.  
 RX MEDLINE=97475217; PubMed=9334741;  
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,  
 RA Sax M.;  
 RT "Structure of the receptor binding fragment HC of tetanus  
 neurotoxin.";  
 RL Nat. Struct. Biol. 4:788-792(1997).  
 CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-1-PHE-77  
 BOND OF SYNAPTOSOMAL-2.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-1-Phe-77 bond in  
 synaptobrevin 2.  
 CC -1- COPACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO  
 YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE  
 AND ARE NON-TOXIC AFTER SEPARATION.  
 CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO  
 GANGLIOSIDE RECEPTORS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X04436; CAA28033.1; -;  
 DR EMBL; X06214; CAA29564.1; -;  
 DR EMBL; AF528097; NAO37454.1; -;  
 DR EMBL; M12739; AAA23282.1; -;  
 DR PIR; A25689; BTCLTN.  
 DR PDB; 1AF9; 29-APR-98.  
 DR PDB; 1A8D; 14-OCT-98.  
 DR PDB; 1D0H; 27-MAR-00.  
 DR PDB; 1DFQ; 24-MAR-00.  
 DR PDB; 1DIW; 24-MAR-00.  
 DR PDB; 1DLL; 24-MAR-00.  
 DR PDB; 1FV3; 05-SEP-01.  
 DR MEROPS; M27.001; -;  
 DR InterPro; IPR000395; Bontoxilysin.  
 DR InterPro; IPR006025; Zn\_Mrpeptidse.  
 DR Pfam; PF01742; Peptidase\_M27; 1.

DR PRINTS; PR00760; BONTOXILYSIN.  
 DR ProDom; PD001963; Bontoxilysin; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;  
 FT 3D-structure; Complete proteome.  
 FT INT\_MET 0 0  
 FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.  
 FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.  
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 233 233 BY SIMILARITY.  
 FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT TRANSMEM 226 246 POTENTIAL.  
 FT TRANSMEM 669 689 POTENTIAL.  
 FT DISULFID 438 466 INTERCHAIN.  
 FT DISULFID 1076 1092  
 FT HELIX 876 882  
 FT TURN 883 883  
 FT STRAND 884 891  
 FT TURN 892 893  
 FT STRAND 894 897  
 FT STRAND 904 907  
 FT TURN 909 910  
 FT STRAND 912 915  
 FT STRAND 920 925  
 FT TURN 928 929  
 FT STRAND 932 935  
 FT HELIX 938 940  
 FT TURN 941 946  
 FT STRAND 949 956  
 FT HELIX 962 968  
 FT TURN 969 970  
 FT STRAND 972 977  
 FT STRAND 980 981  
 FT HELIX 983 985  
 FT STRAND 987 995  
 FT TURN 996 997  
 FT STRAND 998 1004  
 FT TURN 1006 1007  
 FT STRAND 1010 1016  
 FT STRAND 1020 1020  
 FT TURN 1021 1022  
 FT STRAND 1031 1037  
 FT TURN 1039 1040  
 FT STRAND 1042 1047  
 FT TURN 1048 1049  
 FT STRAND 1050 1056  
 FT TURN 1058 1059  
 FT STRAND 1068 1074  
 FT TURN 1079 1080  
 FT STRAND 1082 1091  
 FT HELIX 1097 1105  
 FT TURN 1106 1107  
 FT STRAND 1112 1112  
 FT STRAND 1114 1114  
 FT TURN 1116 1117  
 FT STRAND 1120 1120  
 FT STRAND 1122 1122  
 FT TURN 1123 1124  
 FT STRAND 1127 1131  
 FT HELIX 1132 1134  
 FT TURN 1135 1136  
 FT STRAND 1137 1141  
 FT TURN 1144 1145  
 FT STRAND 1148 1152  
 FT STRAND 1155 1158  
 FT TURN 1159 1162  
 FT STRAND 1163 1166  
 FT STRAND 1173 1178  
 FT TURN 1184 1185  
 FT STRAND 1188 1188  
 FT STRAND 1190 1190  
 FT TURN 1191 1192  
 FT STRAND 1193 1201

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:00:31 ; Search time 6.86704 Seconds  
(without alignments)  
253.383 Million cell updates/sec

Title: US-09-848-834A-14

Perfect score: 200

Sequence: 1 XHWSYGLRPGSSGSLFNFTVFWLRPKVSASHLE 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	114	57.0	1314	1 TETX_CLOTE	P04958 clostridium
2	67	33.5	1296	1 BXG_CLOBO	Q60393 clostridium
3	65	32.5	1295	1 BXA1_CLOBO	P10845 clostridium
4	63	31.5	1274	1 BXF_CLOBO	P30996 clostridium
5	63	31.5	1290	1 BXB_CLOBO	P10844 clostridium
6	62	31.0	1051	1 BPA2_AHSV6	O71024 african hor
7	62	31.0	1295	1 BX22_CLOBO	Q43894 clostridium
8	60	30.0	92	1 GON1_TUPGB	Q95335 tupaia gliis
9	59	29.5	67	1 GON1_MACMU	P55247 macaca mula
10	59	29.5	92	1 GON1_HUMAN	P01148 homo sapien
11	58	29.0	61	1 GON1_SHEEP	Q28588 ovis aries
12	58	29.0	63	1 GON1_MESAU	O09163 mesocricetu
13	58	29.0	89	1 GON1_XENLA	P43656 xenopus lae
14	58	29.0	90	1 GON1_MOUSE	P13562 mus musculus
15	58	29.0	91	1 GON1_PIG	P49921 sus scrofa
16	58	29.0	92	1 GON1_RAT	P07490 rattus norv
17	57.5	28.7	352	1 COA2_SV40	P03093 simian viru
18	56.5	28.2	94	1 GON1_HAPBU	P51918 haplochromi
19	56	28.0	92	1 GON1_CHICK	P37042 gallus gall
20	56	28.0	1250	1 BXE_CLOBO	Q00496 clostridium
21	56	28.0	1250	1 BXE_CLOBO	P30995 clostridium
22	56	28.0	1290	1 BXC1_CLOBO	P18640 clostridium
23	54.5	27.3	464	1 VNSS_TSWV1	P26002 tomato spot
24	54	27.0	10	1 GON1_ALLMI	P37041 alligator m
25	53	26.5	99	1 GON1_DICLA	Q91a10 dicentrarch
26	53	26.5	449	1 VNSS_INSYN	Q01811 impatiens n
27	53	26.5	1196	1 BXCN_CLOBO	P46081 clostridium
28	52	26.0	95	1 GON1_MORGA	O73812 morone saxa
29	52	26.0	95	1 GON1_PAGNA	P70074 pegrus majo
30	52	26.0	95	1 GON1_SPAAU	P51919 sparus aura
31	52	26.0	467	1 VNSS_TSWV1	P26003 tomato spot
32	51.5	25.8	1539	1 SMCY_HUMAN	Q9by66 homo sapien
33	51.5	25.8	1548	1 SMCY_MOUSE	Q62240 mus musculus

34	51.5	25.8	1554	1 SMCX_MOUSE	P41230 mus musculu
35	51.5	25.8	1560	1 SMCX_HUMAN	P41229 homo sapien
36	51	25.5	227	1 RHON_MOUSE	P52198 homo sapien
37	51	25.5	227	1 RHON_MOUSE	Q9qym5 mus musculu
38	51	25.5	956	1 CLS3_HUMAN	Q9bqt9 homo sapien
39	50.5	25.2	1276	1 BXD_CLOBO	P19321 clostridium
40	50	25.0	90	1 GON8_RANDY	Q9iaa2 rana dybows
41	50	25.0	144	1 CNI_DROME	P49858 drosophila
42	50	25.0	585	1 YH70_SVNY3	P73627 synchocyst
43	49.5	24.8	344	1 COA2_POVJC	P03095 polyomaviru
44	49.5	24.8	1115	1 ITA3_DROME	O44386 drosophila
45	49	24.5	92	1 GON1_CAVPO	O54713 cavia porce

ALIGNMENTS

RESULT 1  
TETX\_CLOTE  
ID TETX\_CLOTE STANDARD; PRT; 1314 AA.  
AC P04958:  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains:  
DE Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy  
DE chain (Tetanus toxin chain H)].  
GN TETX OR CTF60.  
OS Clostridium tetani.  
OG Plasmid pE88, and Plasmid 75 Kbp.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC PLASMID=75 Kbp;  
RX MEDLINE=87053814; PubMed=3536478;  
RA Eisel U., Jarasch W., Goretzki K., Henschen A., Engels J.,  
RA Weller U., Hudel M., Habermann E., Niemann H.;  
RT "Tetanus toxin: primary structure, expression in E. coli, and  
RT homology with botulinum toxins.";  
EMBO J. 5:2495-2502(1986).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CN3911; PLASMID=75 Kbp;  
RX MEDLINE=87040747; PubMed=3774547;  
RA Fairweather N.F., Lyness V.A.;  
RT "The complete nucleotide sequence of tetanus toxin.";  
Nucleic Acids Res. 14:7809-7812(1986).  
[3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Massachusetts / E88; PLASMID=pE88;  
RX MEDLINE=22457253; PubMed=12552129;  
RA Brueggemann H., Baeumer S., Fricke W.F., Wiese A., Liesegang H.,  
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,  
RA Gottschalk G.;  
RT "The genome sequence of Clostridium tetani, the causative agent of  
RT tetanus disease.";  
Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).  
[4]  
RP SEQUENCE OF 742-1314 FROM N.A.  
RC PLASMID=75 Kbp;  
RX MEDLINE=86085672; PubMed=3510187;  
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;  
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin  
RT fragment C in Escherichia coli.";  
J. Bacteriol. 165:21-27(1986).  
[5]  
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.  
RX MEDLINE=90201034; PubMed=2108021;  
RA Krieglstein K., Henschen A., Weller U., Habermann E.;  
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups  
RT in tetanus toxin.";

## RESULT 14

RHSHG  
gonadoliberin - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 18-Mar-1997  
C:Accession: A93780; A01411  
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.; Bl  
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972  
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing facto  
A:Reference number: A93780; MUID:72094314; PMID:4550508  
A:Accession: A93780  
A:Molecule type: protein  
A:Residues: 1-10 <BUR>  
A:Note: the natural and synthetic hormones have the same biological activity  
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 29.0%; Score 58; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.041;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10  
|||||  
Db 2 HWSYGLRPG 10

## RESULT 15

151423  
gonadoliberin precursor - African clawed frog  
N:Alternate names: luteinizing hormone releasing hormone  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
C:Accession: I51423  
R:Hayes, W.P.; Wray, S.; Battey, J.F.  
Endocrinology 134, 1835-1845, 1994  
A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved doma  
A:Reference number: I51423; MUID:94185563; PMID:8137750  
A:Accession: I51423  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-89 <HAY>  
A:Cross-references: GB:L28040; NID:g496291; PIDN:AAA49728.1; PID:g496292  
C:Genetics:  
A:Gene: GnRH-I  
C:Superfamily: gonadoliberin

Query Match 29.0%; Score 58; DB 2; Length 89;  
Best Local Similarity 100.0%; Pred. No. 0.49;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10  
|||||  
Db 25 HWSYGLRPG 33

Search completed: September 24, 2003, 17:49:33  
Job time : 14.0166 secs

Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 17 FNNFTVSWLRYPK 30  
Db 297 YQNFSEWVRIPK 310

## RESULT 11

I78541  
gonadoliberin precursor - rhesus macaque (fragment)  
N:Alternate names: luteinizing hormone releasing hormone  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999  
C:Accession: I78541  
R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.  
Neuroendocrinology 60, 346-359, 1994  
A:Title: Developmental expression of the genes encoding transforming growth factor alpha  
A:Reference number: I58134; MUID:95124501; PMID:7545971  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Accession: I78541  
A:Molecule type: mRNA  
A:Residues: 1-67 <RES>  
A:Cross-references: GB:S75918; NID:9912831; PIDN:AAH33096.1; PID:g912832  
C:Superfamily: gonadoliberin

Query Match 29.5%; Score 59; DB 2; Length 67;

Best Local Similarity 52.4%; Pred. No. 0.25; Mismatches 3; Indels 5; Gaps 1;

QY 2 HWSYGLRPGS--SGPSLFNNF 20  
Db 7 HWSYGLRPGGKRDAENLMDSF 27

## RESULT 12

RHHUG  
gonadoliberin precursor [validated] - human  
N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing  
N:Contains: gonadoliberin-associated protein (GAP); progonadoliberin  
C:Species: Homo sapiens (man)  
C:Date: 17-Mar-1987 #sequence\_revision 21-Jul-1995 #text\_change 08-Dec-2000  
C:Accession: S05308; A26173; A93442; A90108; A01410; S45718  
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.  
Nucleic Acids Res. 17, 6403-6404, 1989  
A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone gene  
A:Reference number: S05308; MUID:89366682; PMID:2671939  
A:Accession: S05308

A:Status: translation not shown

A:Molecule type: DNA  
A:Residues: 1-92 <HAY>  
A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956  
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot  
A:Reference number: A94090; MUID:86094338; PMID:2867548  
A:Accession: A26173

A:Molecule type: mRNA

A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749  
A:Experimental source: hypothalamus  
R:Seeburg, P.H.; Adelman, J.P.  
Nature 311, 666-668, 1984

A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasing h  
A:Reference number: A93342; MUID:85012739; PMID:6090951  
A:Accession: A93342

A:Molecule type: mRNA

A:Residues: 1-15, 'S', 17-92 <SEE>  
A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357  
A:Experimental source: placenta  
R:Rian, L.; Rousseau, P.  
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982

A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in  
A:Reference number: A90108; MUID:83126573; PMID:6760865

A:Accession: A90108  
A:Molecule type: protein  
A:Residues: 24-33 <IAMS>  
A:Experimental source: placental trophoblasts  
R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amara  
FEBS Lett. 346, 203-206, 1994  
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone  
A:Reference number: S45718; MUID:94283597; PMID:8013634  
A:Contents: annotation; degradation pathway of synthetic hormone

C:Genetics:

A:Gene: GDB:GNRH; LHRH; GRH

A:Cross-references: GDB:1333746; OMIM:227200; OMIM:152760

A:Map position: 8p21-8p11.2

A:Introns: 47/3; 79/3

C:Function:

A:Description: gonadoliberin stimulates pituitary secretion of lutropin and follitrop  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting acti  
C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-92/Product: progadoliberin #status predicted <PGN>  
F:24-33/Product: gonadoliberin #status experimental <MAT>  
F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 29.5%; Score 59; DB 1; Length 92;

Best Local Similarity 52.4%; Pred. No. 0.36;

Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 2 HWSYGLRPGS--SGPSLFNNF 20  
Db 25 HWSYGLRPGGKRDAENLIDSF 45

## RESULT 13

RHPGG

gonadoliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 18-Mar-1997

C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90172; MUID:72114303; PMID:4946067

A:Accession: A01411

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A:Reference number: A90176; MUID:72065376; PMID:4942726

A:Contents: annotation; synthesis

A:Note: the synthetic and natural A.V.

R:Baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544; PMID:4946275

A:Contents: annotation

A:Note: Trp-3 appears to be essential for biological activity

C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing an

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 29.0%; Score 58; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.041;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

Db 2 HWSYGLRPG 10

R;Schivo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta Nature 359, 832-835, 1992

A>Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic A:Reference number: S27125; MUID:93063293; PMID:1331807

C:Contents: annotation

C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic s C:Genetics:

A:Gene: bont/b

C:Function:

A>Description: catalyzes hydrolysis of a GlN-Phe peptide bond in synaptobrevin 2 C:Superfamily: tetanus toxin

C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc F;2-441/Product: botoxilysin B light chain #status experimental <LGH T>

F;442-1291/Product: botoxilysin B heavy chain #status experimental <HVT>

F;230\_234/Binding site: zinc (His) #status predicted

F;231/Active site: Gl u #status predicted

Query Match 31.5%; Score 63; DB 1; Length 1291;  
Best Local Similarity 62.5%; Pred. No. 2;  
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30  
I:: I:: I:: I:: I:: I::

D b 921 SVFLDFSVMIRIPK 936

RESULT 9

I40645

botulinum neurotoxin type A - Clostridium botulinum

C:Species: Clostridium botulinum

C>Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 16-Jul-1999

C:Accession: I40645

R;Williams, A.; East, A.K.; Lawson, P.A.; Collins, M.D.  
Res. Microbiol. 144, 547-556, 1993

A>Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum typ A:Reference number: I40645; MUID:94143603; PMID:8310180

A:Accession: I40645

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1296 <RES>

A:Cross-references: EMBL:X73423; NID:g507070; PIDN:CAA51824.1; PID:g507071

C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

Query Match 31.0%; Score 62; DB 2; Length 1296;  
Best Local Similarity 50.0%; Pred. No. 2.7;  
Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30  
I:: I:: I:: I:: I:: I::

D b 936 SMYENFTSFWIKIPK 951

RESULT 10

S48110

neurotoxin type F - Clostridium botulinum (fragment)

C:Species: Clostridium botulinum

C>Date: 14-Jul-1995 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999

C:Accession: S48110

R;Campbell, K.D.; Collings, M.D.; East, A.K.  
J. Clin. Microbiol. 31, 2255-2262, 1993

A>Title: Gene probes for identification of the botulinus neurotoxin gene and specific A:Reference number: S48103; MUID:94013372; PMID:8408542

A:Accession: S48110

A>Status: preliminary; translation not shown

A:Molecule type: DNA

A:Residues: 1-366 <CAM>

A:Cross-references: EMBL:X70821; NID:g407792; PIDN:CAA50152.1; PID:g407793

C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

Query Match 30.5%; Score 61; DB 2; Length 366;  
Best Local Similarity 57.1%; Pred. No. 0.91;

QY 15 SLFNNFTVSFWLRVPK 30  
I : : : : :  
Db 936 SMYENFSTSFWRIRPK 951

## RESULT 4

S33411  
botulinum neurotoxin type F - Clostridium baratii  
C:Species: Clostridium baratii  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 16-Jul-1999  
C:Accession: S33411; S31860  
R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T.  
FEMS Microbiol. Lett. 108, 175-182, 1993  
A:Title: Nucleotide sequence of the gene coding for Clostridium baratii type F neurotoxin  
A:Reference number: S33411; MUID:93252228; PMID:8486245  
A:Accession: S33411  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1268 <THO>  
A:Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139  
C:Superfamily: tetanus toxin  
C:Keywords: neurotoxin

Query Match 32.0%; Score 64; DB 2; Length 1268;  
Best Local Similarity 62.5%; Pred. No. 1.4;  
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30  
I : : : : :  
Db 920 SRYQNFVSFWVRIPK 935

## RESULT 5

I40631  
non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum  
C:Species: Clostridium botulinum  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 16-Jul-1999  
C:Accession: I40631; S48103; S48104; S36015  
R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.  
Curr. Microbiol. 28, 101-110, 1994  
A:Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulinum  
A:Reference number: I40631; MUID:94122659; PMID:7764370  
A:Accession: I40631  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1291 <RES>  
A:Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149  
R:Campbell, K.D.; Collins, M.D.; East, A.K.  
J. Clin. Microbiol. 31, 2255-2262, 1993  
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific isoforms

A:Reference number: S48103; MUID:94013372; PMID:8408542  
A:Accession: S48103  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 634-761, 'b', 763-841, 'M', 843, 'T', 845, 'N', 847-994 <CAMI>  
A:Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779  
A:Experimental source: non-proteolytic strain 2129B (Scott)  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993  
A:Accession: S48104  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 634-843, 'T', 845, 'N', 847-994 <CAM>  
A:Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781  
A:Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)

C:Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic site  
C:Genetics:  
A:Gene: bont/b  
C:Superfamily: tetanus toxin  
C:Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc  
F:2-441/Product: botulinum neurotoxin type B light chain #status predicted <LIGHT>  
F:442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVT>  
F:230, 234/Binding site: zinc (His) #status predicted

F:231/Active site: Glu #status predicted

Query Match 32.0%; Score 64; DB 2; Length 1291;  
Best Local Similarity 62.5%; Pred. No. 1.4;  
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30  
I : : : : :  
Db 921 SMFLDFSFWIRIPK 936

## RESULT 6

S48109  
neurotoxin type F - Clostridium botulinum (fragment)  
C:Species: Clostridium botulinum  
C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 16-Jul-1999  
C:Accession: S48109  
R:Campbell, K.D.; Collins, M.D.; East, A.K.  
J. Clin. Microbiol. 31, 2255-2262, 1993  
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific isoforms  
A:Reference number: S48103; MUID:94013372; PMID:8408542  
A:Accession: S48109  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-369 <CAM>  
A:Cross-references: EMBL:X70820; NID:g407790; PIDN:CAA50151.1; PID:g407791  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993  
C:Superfamily: tetanus toxin

Query Match 31.5%; Score 63; DB 2; Length 369;  
Best Local Similarity 56.2%; Pred. No. 0.48;  
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30  
I : : : : :  
Db 295 SRYQNFVSFWVRIPK 310

## RESULT 7

I40813  
neurotoxin type F - Clostridium botulinum  
C:Species: Clostridium botulinum  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 16-Jul-1999  
C:Accession: I40813; S48108  
R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, D.E.  
FEMS Microbiol. Lett. 96, 225-230, 1992  
A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum  
A:Reference number: I40644  
A:Accession: I40813  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1274 <RES>  
A:Cross-references: GB:M2906; NID:g144866; PIDN:AAA23263.1; PID:g144867  
R:Campbell, K.D.; Collins, M.D.; East, A.K.  
J. Clin. Microbiol. 31, 2255-2262, 1993  
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific isoforms  
A:Reference number: S48103; MUID:94013372; PMID:8408542  
A:Accession: S48108  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 634-1002 <CAM>  
A:Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789  
C:Superfamily: tetanus toxin  
C:Keywords: neurotoxin

Query Match 31.5%; Score 63; DB 2; Length 1274;  
Best Local Similarity 56.2%; Pred. No. 1.9;  
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30  
I : : : : :  
Db 928 SRYQNFVSFWVRIPK 943



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:30:06 ; Search time 13.0166 Seconds  
(without alignments)  
273.361 Million cell updates/sec

Title: US-09-848-834A-14

Perfect score: 200

Sequence: 1 XHWSYGLRPGSSGSLFNNFTVSWLRVPRKVSASHLE 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	57.0	1315	1 BTCLTN	tentoxilysin (EC 3
2	67	33.5	1297	2 S39791	neurotoxin - Clost
3	65	32.5	1296	1 BTCLAB	bontoxilysin (EC 3
4	64	32.0	1268	2 S3411	botulinum neurotox
5	64	32.0	1291	2 I40631	non-proteolytic bo
6	63	31.5	369	2 S48109	neurotoxin type F
7	63	31.5	1274	2 I40813	neurotoxin type F
8	63	31.5	1291	1 A48940	bontoxilysin (EC 3
9	62	31.0	1296	2 I40645	botulinum neurotox
10	61	30.5	366	2 S48110	neurotoxin type F
11	59	29.5	67	2 I78541	gonadoliblerin prec
12	59	29.5	92	1 RHUG	gonadoliblerin prec
13	58	29.0	10	1 RHSG	gonadoliblerin - pi
14	58	29.0	10	1 RHSG	gonadoliblerin - sh
15	58	29.0	89	2 I51423	gonadoliblerin prec
16	58	29.0	90	1 RHSG	gonadoliblerin prec
17	58	29.0	92	1 RHSG	gonadoliblerin prec
18	57.5	28.7	352	1 VVVP24	coat protein VP2 -
19	56.5	28.2	98	2 I50739	gonadotropin-relea
20	56	28.0	92	2 I50644	gonadoliblerin I pr
21	56	28.0	367	2 S48106	neurotoxin type E
22	56	28.0	1251	2 JH0256	botulinum neurotox
23	56	28.0	1252	2 S21178	botulinum neurotox
24	56	28.0	1291	2 S46431	botulinum neurotox
25	56	28.0	1291	2 A49777	botulinum neurotox
26	54.5	27.3	91	2 JC7393	medaka-type gonado
27	54.5	27.3	464	1 MNVWC	nonstructural prot
28	54	27.0	10	1 RHQ1	gonadoliblerin I -
29	53.5	26.8	251	2 AD1669	3'-exo-deoxyribonu

#### ALIGNMENTS

##### RESULT 1

###### BTCLTN

tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani

N:Alternate names: tetanus neurotoxin

C:Species: Clostridium tetani

C>Date: 31-Mar-1988 #sequence-revision 31-Mar-1988 #text\_change 03-Jun-2002

C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364

R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hucl, E.

EMBO J. 5, 2495-2502, 1986

A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with

A:Reference number: A25689; MUID:87053814; PMID:87053814

A:Accession: A25689

A:Molecule type: DNA

A:Residues: 1-1315 <EIS>

A:Cross-references: GB:X04436; NID:940769; PIDN:CAA28033.1; PID:940770

R:Fairweather, N.F.; Lyness, V.A.

Nucleic Acids Res. 14, 7809-7812, 1986

A:Title: The complete nucleotide sequence of tetanus toxin.

A:Reference number: A25757; MUID:87040747; PMID:3774547

A:Accession: A25757

A:Molecule type: DNA

A:Residues: 1-1315 <FAI>

A:Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:940774

R:Experimental source: strain CN3911

R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.

J. Bacteriol. 165, 21-27, 1986

A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment :

A:Reference number: A25194; MUID:86085672; PMID:3510187

A:Accession: A25194

A:Molecule type: DNA

A:Residues: 743-1315 <FA2>

A:Cross-references: GB:M12739; NID:9144920; PIDN:AAA23282.1; PID:9144921

A:Accession: B25194

A:Molecule type: protein

A:Residues: 865-894 <FA3>

R:Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.

Infect. Immun. 57, 3588-3593, 1989

A:Title: Isolation, purification, and characterization of fragment B, the NH-2-termi

A:Reference number: A60759; MUID:90035436; PMID:2478476

A:Accession: A60759

A:Molecule type: protein

A:Residues: 461-475 <MAT>

R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.

J. Immunol. 142, 394-402, 1989

A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.

A:Reference number: JS0098; MUID:89093918; PMID:2463305

R:Contents: annotation; epitope region

R:Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,

Nature 359, 832-835, 1992

A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo

A:Reference number: S27125; MUID:93063293; PMID:1331807

A:Contents: annotation

DR EMBL; AF156107; AAD38991.1; -.  
DR EMBL; AF156105; AAD38944.1; -.  
DR InterPro; IPR001070; Polyoma\_coat2.  
DR Pfam; PF00761; Polyoma\_coat2; 1.  
KW Coat protein.  
SQ SEQUENCE 234 AA; 26962 MW; 2490

Search completed: September 24, 2003, 17:47:19  
Job time : 35.8726 secs

```
RESULT 13
Q90Y63          PRELIMINARY;      PRT;      90 AA.
AC Q90Y63;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I
DE (GAP1)].
GN GnRH I OR GnRH.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX TISSUE=Forebrain;
RX MEDLINE=21102951; PubMed=11170016;
RA Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA Kwop H.B.;
RT "Cloning and characterization of cDNAs encoding the GnRH1 and GnRH2
RT precursors from bullfrog (Rana catesbeiana).";
RL J. Exp. Zool. 289:190-201(2001).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- TISSUE SPECIFICITY: FOREBRAIN.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT SIGNIFICANTLY HIGHER LEVELS
CC DURING POST-BREEDING. NOT EXPRESSED IN PITUITARY.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR EMBL; AF188754; AAL05972.1; -.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRNI.
DR PROSITE; PS00473; GnRH; 1.
DR Cleavage on pair of basic residues; Hormone; Amidation; Signal;
KW Pyroglutamate carboxylic acid.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 90 PROGONADOLIBERIN I.
FT PEPTIDE 25 34 GONADOLIBERIN I.
FT PEPTIDE 38 86 GnRH-ASSOCIATED PEPTIDE I
FT MOD_RES 25 25 PYROGLUTAMINE CARBOXYLIC ACID
FT MOD_RES 34 34 AMIDATION (G-35 PROVIDE AMIDE GROUP)
FT SEQUENCE 90 AA; 10291 MW; 317203B4E3DA2FE7 CRC64;
Query Match 29.0%; Score 58; DB 13; Length 90;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 26 HWSYGLRPG 34

RESULT 14
Q9PRH0          PRELIMINARY;      PRT;      91 AA.
AC Q9PRH0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Prepro-mgNRH precursor (Gonadoliberin) (Gonadotropin-releasing
DE hormone) (LH-RH) (Luliberin).
OS Anquilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguilla.
```

```
OX NCBI_TaxID=7937;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Okubo K., Suetake H., Aida K.;
RT "Expression of two gonadotropin-releasing hormone (GnRH) precursor
RT genes in various tissues of the Japanese eel and evolution of GnRH.";
RL Zool. Sci. 16:471-478(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Okubo K., Suetake H., Aida K.;
RT "A splicing variant for the prepro-mammalian gonadotropin-releasing
RT hormone (prepro-mgNRH) mRNA is present in the brain and various
RT peripheral tissues of the Japanese eel.";
RL Zool. Sci. 16:645-651(1999).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR EMBL; AB026989; BAA82608.1; -.
DR EMBL; AB026991; BAA83597.1; -.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRNI.
DR PROSITE; PS00473; GnRH; 1.
KW Amidation; Hormone; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 32 MGNRH.
FT CHAIN 33 91 GnRH ASSOCIATED PEPTIDE.
FT SEQUENCE 91 AA; 9893 MW; BA15C9DC08434AYB CRC64;
Query Match 29.0%; Score 58; DB 13; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 24 HWSYGLRPG 32

RESULT 15
Q9W9A5          PRELIMINARY;      PRT;      234 AA.
AC Q9W9A5;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE Viral coat protein 3.
OS Simian virus 40 (SV40).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VA45-54-1, and VA45-54-2;
RX MEDLINE=97368433; PubMed=9225047;
RA Lednický J.A., Butel J.S.;
RT "Tissue culture adaptation of natural isolates of simian virus 40:
RT changes occur in viral regulatory region but not in carboxy-terminal
RT domain of large T-antigen.";
RL J. Gen. Virol. 78:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VA45-54-1, and VA45-54-2;
RA Lednický J.A., Stewart A.R., Butel J.S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=VA45-54-2; PubMed=9584955;
RX MEDLINE=98244346;
RA Stewart A.R., Lednický J.A., Butel J.S.;
RT "Sequence analyses of human tumor-associated SV40 DNAs and SV40 viral
RT isolates from monkeys and humans.";
RL J. Neurovirol. 4:182-193(1998).
```

SQ SEQUENCE 1291 AA; 150574 MW; 0227CAEF4F58504D



```

RESULT 3
QID Q45848 PRELIMINARY; PRT; 361 AA.
ID Q45848;
AC Q45848;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Botulinum neurotoxin type B (Fragment).
DE BONT/B.
GN
GN
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC Clostridium.
OX NCBI_TaxID=1491;
OX [1]
RN
RN
RP SEQUENCE FROM N.A.
RC STRAIN=type B;
RX MEDLINE=94013372; PubMed=8408542;
RT Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RL specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL; X70819; CAA50150.1; -
DR HSSP; P10845; 3BTA.
KW Neurotoxin.
FT
FT NON_TER 1
FT NON_TER 361
FT FT
SQ SEQUENCE 361 AA; 42131 MW; A2E0FFFC81F9533D CRC64;

Query Match 32.0%; Score 64; DB 2; Length 361;
Best Local Similarity 62.5%; Pred. No. 0.77;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 15 SLFNFTVSWFLRVPK 30
I:|:|:|:|:|:|
Db 288 SMFLDFSVSWIRPK 303

```

ID	Q45846	PRELIMINARY;	PRT;	361 AA.
DT	Q45846;			
DT	01-NOV-1996 (Tremblrel. 01, Created)			
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)			
DT	01-OCT-2000 (Tremblrel. 15, Last annotation update)			
DE	Botulinum neurotoxin type B (Fragment).			
GN	BONT/B.			
OS	Clostridium botulinum.			
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=type B;			
RX	MEDLINE=94013372; PubMed=8408542;			
RT	Campbell K., East A.K., Collins M.D.;			
RT	*Gene probes for identification of the botulin neurotoxin gene and			
RL	specific identification of neurotoxin types B, E, and F.*;			
RL	J. Clin. Microbiol. 31:2255-2262(1993).			
DR	EMBL; X70814; CAA50145.1; -			
DR	HSSP; P10845; 3BTA.			
KW	Neurotoxin.			
FT	NON_TER	1		
FT	NON_TER	361		
FT	NON_TER	361		
SQ	SEQUENCE	361 AA;	42175 MW; 533EA98735CD98E1 CRC64;	
	Query Match	32.0%;	Score 64; DB 2; Length 361;	
	Best Local Similarity	62.5%;	Pred. No. 0.77;	
	Matches	10; Conservative	5; Mismatches	1; Indels
				0; Gaps
QY	15 SLENNFTVSFWLRPK 30			
	:   : :			
Db	288 SMELDFSVFWIRPK 303			
RESULT 5				
Q9X708				
ID	Q9X708	PRELIMINARY;	PRT;	441 AA.
AC	Q9X708;			
DT	01-NOV-1999 (Tremblrel. 12, Created)			
DT	01-NOV-1999 (Tremblrel. 12, Last sequence update)			
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)			
DE	Botulinum neurotoxin type B (fragment).			
GN	BONT/B.			
OS	Clostridium botulinum.			
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=99343691; PubMed=10413679;			
RX	Lalli G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,			
RA	Schiavo G.;			
RT	*Functional characterisation of tetanus and botulinum neurotoxins			
RT	binding domains.*;			
RL	J. Cell Sci. 112:2715-2724(1999).			
DR	EMBL; AJ242628; CAB43706.1; -			
DR	HSSP; P10845; 3BTA.			
KW	Neurotoxin.			
FT	NON_TER	1		
FT	NON_TER	441		
FT	NON_TER	441		
SQ	SEQUENCE	441 AA;	52772 MW; 721DOB468E8C95A4 CRC64;	
	Query Match	32.0%;	Score 64; DB 2; Length 441;	
	Best Local Similarity	62.5%;	Pred. No. 0.96;	
	Matches	10; Conservative	5; Mismatches	1; Indels
				0; Gaps
QY	15 SLENNFTVSFWLRPK 30			
	:   : :			
Db	79 SMELDFSVFWIRPK 94			
RESULT 6				

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:29:11 ; Search time 35.8726 Seconds  
(without alignments)  
266.163 Million cell updates/sec

Title: US-09-848-834A-14  
Perfect score: 200  
Sequence: 1 XHWSYGLRPGSSGSLFNFTVSEWLRVPKVSASHLE 37

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	57.0	451	Q9LA13	Q9LA13 clostridium
2	114	57.0	1310	Q93N27	Q93N27 clostridium
3	64	32.0	361	Q45848	Q45848 clostridium
4	64	32.0	361	Q45846	Q45846 clostridium
5	64	32.0	441	Q9X708	Q9X708 clostridium
6	64	32.0	1268	Q45851	Q45851 clostridium
7	64	32.0	1291	Q92AJ8	Q92AJ8 clostridium
8	64	32.0	1291	Q93G71	Q93G71 clostridium
9	64	32.0	1291	Q933K0	Q933K0 clostridium
10	64	32.0	1291	Q08077	Q08077 clostridium
11	64	32.0	1291	Q08GR6	Q08GR6 clostridium
12	61	30.5	1278	Q57236	Q57236 clostridium
13	58	29.0	90	Q90Y63	Q90Y63 rana catesb
14	58	29.0	91	Q9PRH0	Q9PRH0 anguilla ja
15	57.5	28.7	234	Q9W9A5	Q9W9A5 simian viru
16	57.5	28.7	234	Q92837	Q92837 simian viru

17	57.5	28.7	352	12	Q98VM1	Q98vm1 simian viru
18	57.5	28.7	352	12	Q9W9F7	Q9w9f7 simian viru
19	57.5	28.7	352	12	Q910V5	Q910v5 simian viru
20	57	28.5	1280	2	Q92AJ5	Q92aj5 clostridium
21	56.5	28.2	94	13	Q8JFY3	Q8jfy3 oreochromis
22	56	28.0	367	2	Q45861	Q45861 clostridium
23	56	28.0	367	2	Q45862	Q45862 clostridium
24	56	28.0	1251	2	Q9K395	Q9k395 clostridium
25	56	28.0	1252	2	Q8K2M3	Q8k2m3 clostridium
26	56	28.0	1255	2	Q9FAR6	Q9far6 clostridium
27	56	28.0	1291	2	Q93HT3	Q93ht3 clostridium
28	55	27.5	476	5	Q9N3R6	Q9n3r6 caenorhabdi
29	54.5	27.3	91	13	Q8JIO7	Q8jio7 oryzias lat
30	54.5	27.3	91	13	Q9DGC8	Q9dgc8 oryzias lat
31	54.5	27.3	467	12	Q8JXK2	Q8jxk2 tomato spot
32	54.5	27.3	467	12	Q8JVL0	Q8jvl0 tomato spot
33	53.5	26.8	251	16	Q92AM6	Q92am6 listeria in
34	53.5	26.8	609	12	Q8B6X9	Q8b6x9 peste-des-p
35	53	26.5	1196	2	Q45916	Q45916 clostridium
36	53	26.5	1196	2	Q53550	Q53550 clostridium
37	53	26.5	1196	2	Q9LBR2	Q9lbr2 clostridium
38	53	26.5	1196	2	Q9LBS8	Q9lbs8 clostridium
39	53	26.5	1196	2	Q93HT4	Q93ht4 clostridium
40	53	26.5	1196	9	Q92X77	Q92x77 clostridium
41	53	26.5	1196	9	Q38197	Q38197 clostridium
42	53	26.5	1285	2	Q45967	Q45967 clostridium
43	53	26.5	1285	2	Q9LBR1	Q9lbr1 clostridium
44	53	26.5	1702	12	Q8JX15	Q8jx15 norwalk-lik
45	53	26.5	1702	12	Q8JX14	Q8jx14 norwalk-lik

## ALIGNMENTS

## RESULT 1

Q9LA13 PRELIMINARY; PRT; 451 AA.  
AC Q9LA13;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Tetanus toxin (Fragment).  
OS Clostridium tetani.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=20886;  
RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;  
RT "Fragment C of Tetanus Toxin."  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF154828; AAF73267.1; -.  
DR HSSP; P04958; IABD.  
DR InterPro; IPR001064; Crystallin.  
DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
FT NON\_TER 1  
SQ SEQUENCE 451 AA; 51823 MW; 69A8C5F030E6CD8E CRC64;

Query Match 57.0%; Score 114; DB 2; Length 451;  
Best Local Similarity 95.5%; Pred. No. 7e-08;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37

Db 82 MENNFTVSFWLRVPKVSASHLE 103

## RESULT 2

Q93N27 PRELIMINARY; PRT; 1310 AA.  
ID Q93N27  
AC Q93N27;  
DT 01-DEC-2001 (Tremblrel. 19, Created)

; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913.880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 16  
; LENGTH: 858  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 458 to 1315 of SEQ ID NO: 1  
US-08-913-880C-16

Query Match 46.9%; Score 76; DB 4; Length 858;  
Best Local Similarity 93.8%; Pred. No. 0.005;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 16 LOYIKANSKFIGITEL 31  
Db 372 MQYIKANSKFIGITEL 387

RESULT 14  
US-08-913-880C-15  
; Sequence 15, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913.880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 15  
; LENGTH: 860  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 456 to 1315 of SEQ ID NO: 1  
US-08-913-880C-15

Query Match 46.9%; Score 76; DB 4; Length 860;  
Best Local Similarity 93.8%; Pred. No. 0.005;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 16 LOYIKANSKFIGITEL 31  
Db 374 MQYIKANSKFIGITEL 389

RESULT 15  
US-08-913-880C-14  
; Sequence 14, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913.880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 14  
; LENGTH: 862  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 454 to 1315 of SEQ ID NO: 1  
US-08-913-880C-14

Query Match 46.9%; Score 76; DB 4; Length 862;  
Best Local Similarity 93.8%; Pred. No. 0.005;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 16 LOYIKANSKFIGITEL 31  
Db 376 MQYIKANSKFIGITEL 391  
Search completed: September 24, 2003, 17:00:24  
Job time : 9.28837 secs

```
; Sequence 35, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-35

Query Match 46.9%; Score 76; DB 1; Length 47;
Best Local Similarity 84.2%; Pred. No. 0.00018;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 GPSLOYIKANSKFIGITEL 31
Db 17 GKKQYIKANSKFIGITEL 35

RESULT 11
US-08-488-351A-35
; Sequence 35, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-35

Query Match 46.9%; Score 76; DB 2; Length 47;
Best Local Similarity 84.2%; Pred. No. 0.00018;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 GPSLOYIKANSKFIGITEL 31
Db 17 GKKQYIKANSKFIGITEL 35

RESULT 12
US-08-913-880C-17
; Sequence 17, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 17
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 463 to 1315 of SEQ ID NO: 1
; US-08-913-880C-17

Query Match 46.9%; Score 76; DB 4; Length 853;
Best Local Similarity 93.8%; Pred. No. 0.0049;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31
Db 367 MOYIKANSKFIGITEL 382

RESULT 13
US-08-913-880C-16
; Sequence 16, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
```

Best Local Similarity 84.2%; pred. No. 0.00014;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

## RESULT 9

```

; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:

```

ZIP: 10134 0033  
COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; OPERATING SYSTEM: PC DOS/MS DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/488,351A  
FILING DATE: 7-JUN-1995

CLASSIFICATION:  
PRIOR APPLICATION

APPLICATION NUMBER: US 08/446,692  
FILING DATE: 7-JUN-1995

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 08/229,275  
FILING DATE: 14-APR-1994

CLASSIFICATION: 424  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/057,166  
FILING DATE: 27-APR-1992  
CLASSIFICATION: A24

CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:

NAME: MALLA C.H. LIA  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4140

REFERENCE/DOCREF NUMBER: 1151 4140 039  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)415-8745

TELEPHONE: (516) 751-6849  
TELEFAX: (516) 751-6849  
INFORMATION FOR SEO ID NO: 63:

```

;
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 37 amino acids
;

```

```

;
; TYPE: amino acid
; TOPOLOGY: linear
;

```

; MOLECULE TYPE: peptide  
US-08-488-351A-63

Query Match	46.9%	Score 76;	DB 2;	Length 37;
-------------	-------	-----------	-------	------------

Best Local Similarity 84.2%; Pred. No. 0.00014;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 GPSLQYIKANSKFIGITEL 31

Db 1 GGKKQYIKANSKFIGITEL 19

RESULT 10

RESULT TO  
US-08-446-692-35

PCT-US95-08596-23

Query Match 46.9%; Score 76; DB 5; Length 17;  
Best Local Similarity 93.8%; Pred. No. 5.7e-05;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31  
:|||||  
Db 2 MOYIKANSKFIGITEL 17

RESULT 5

PCT-US93-11703-63

; Sequence 63, Application PC/TUS9311703

; GENERAL INFORMATION:

; APPLICANT: Chiron Mimotopes Pty. Ltd.

; TITLE OF INVENTION: T-cell Epitopes

; NUMBER OF SEQUENCES: 75

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Grant D. Green

; STREET: 4560 Horton St.

; CITY: Emeryville

; STATE: CA

; COUNTRY: USA

; ZIP: 94608

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30B

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/11703

; FILING DATE: 28-DEC-1993

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/984,852

; FILING DATE: 02-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Green, Grant D.

; REGISTRATION NUMBER: 31,259

; REFERENCE/DOCKET NUMBER: 0222.101

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 510-601-2706

; TELEFAX: 510-655-3542

; INFORMATION FOR SEQ ID NO: 63:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 31 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

PCT-US93-11703-63

Query Match

Best Local Similarity 46.9%; Score 76; DB 5; Length 31;

Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31

:|||||

Db 9 MOYIKANSKFIGITEL 24

RESULT 6

US-08-446-692-57

; Sequence 57, Application US/08446692

; Patent No. 5759551

; GENERAL INFORMATION:

; APPLICANT: Ladd, Anna

; APPLICANT: Wang, Chang Yi

; APPLICANT: Zamb, Timothy

; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maria C.H. Lin

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10154-0053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maria C.H. Lin

; REGISTRATION NUMBER: 29,323

; REFERENCE/DOCKET NUMBER: 1151-4146 US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)415-8745

; TELEFAX: (516)751-6849

; INFORMATION FOR SEQ ID NO: 57:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 37 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-446-692-57

Query Match

Best Local Similarity 46.9%; Score 76; DB 1; Length 37;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 GPSLOYIKANSKFIGITEL 31

:|||||

Db 17 GKKQYIKANSKFIGITEL 35

RESULT 7

US-08-446-692-63

; Sequence 63, Application US/08446692

; Patent No. 5759551

; GENERAL INFORMATION:

; APPLICANT: Ladd, Anna

; APPLICANT: Wang, Chang Yi

; APPLICANT: Zamb, Timothy

; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maria C.H. Lin

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10154-0053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maria C.H. Lin

; REGISTRATION NUMBER: 29,323

; REFERENCE/DOCKET NUMBER: 1151-4146 US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)415-8745

APPLICATION NUMBER: US/08/472,701

Result No.	Score	Query Match	Length	DB	ID	Description
1	84	51.9	24	4	US-08-464-496-31	Sequence 31, Appl
2	84	51.9	24	5	PCT-US92-07218-31	Sequence 31, Appl
3	76	46.9	17	4	US-08-472-701-23	Sequence 23, Appl
4	76	46.9	17	5	PCT-US95-08596-23	Sequence 23, Appl
5	76	46.9	31	5	PCT-US93-11703-63	Sequence 63, Appl
6	76	46.9	37	1	US-08-446-692-57	Sequence 57, Appl
7	76	46.9	37	2	US-08-446-692-63	Sequence 63, Appl
8	76	46.9	37	1	US-08-488-351A-57	Sequence 57, Appl
9	76	46.9	37	2	US-08-488-351A-63	Sequence 63, Appl
10	76	46.9	47	1	US-08-446-692-35	Sequence 35, Appl
11	76	46.9	47	2	US-08-488-351A-55	Sequence 35, Appl
12	76	46.9	853	4	US-08-913-880C-17	Sequence 17, Appl
13	76	46.9	858	4	US-08-913-880C-16	Sequence 16, Appl
14	76	46.9	860	4	US-08-913-880C-15	Sequence 15, Appl
15	76	46.9	862	4	US-08-913-880C-14	Sequence 14, Appl
16	76	46.9	865	4	US-08-913-880C-13	Sequence 13, Appl
17	76	46.9	866	4	US-08-913-880C-12	Sequence 12, Appl
18	76	46.9	874	4	US-08-913-880C-11	Sequence 11, Appl
19	76	46.9	875	4	US-08-913-880C-10	Sequence 10, Appl
20	76	46.9	1315	4	US-08-913-880C-1	Sequence 1, Appl
21	74	45.7	15	2	US-08-319-704-10	Sequence 10, Appl
22	74	45.7	15	2	US-08-661-052-6	Sequence 6, Appl
23	74	45.7	15	2	US-08-460-502-7	Sequence 7, Appl
24	74	45.7	15	3	US-09-046-373-2	Sequence 2, Appl
25	74	45.7	15	3	US-09-188-082-6	Sequence 6, Appl
26	74	45.7	15	4	US-09-364-088-6	Sequence 6, Appl
27	74	45.7	15	4	US-09-102-716-6	Sequence 6, Appl



Query Match 50.0%; Score 81; DB 21; Length 116;  
 Best Local Similarity 76.2%; Pred. No. 0.00045;  
 Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 SSGPSLQYIKANSKFIGITEL 31  
 :| :|||||  
 Db 24 TSNETMQYIKANSKFIGITEL 44

## RESULT 15

AAB45526  
 ID AAB45526 standard; Protein; 116 AA.

XX AC AAB45526;

XX DT 26-FEB-2001 (first entry)

XX Modified murine interleukin-5 SEQ ID NO: 52.

XX Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;  
 KW cancer; eosinophilia; vaccine; allergic rhinitis.

XX OS Mus musculus.

XX OS Clostridium tetani.

XX PN WO200065058-A1.

XX PD 02-NOV-2000.

XX PF 19-APR-2000; 2000WO-DK00205.

XX PR 23-APR-1999; 99DK-0000552.

XX PR 06-MAY-1999; 99US-0132811.

XX PA (MEBI-) M & E BIOTECH AS.

XX PI Klysner S;

XX WPI: 2000-672791/65.

XX N-PSDB; AAC68879.

PT Down-regulating interleukin 5 (IL-5) activity in humans by  
 PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,  
 PT prophylaxis or amelioration of asthma or other chronic allergic  
 PT conditions -

PS Disclosure; Page 159-160; 172pp; English.

XX The present invention is concerned with methods of treating asthma,  
 CC eosinophilia, allergic rhinitis and other allergic diseases. These  
 CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5  
 CC proteins and their coding sequences to down-regulate IL-5 activity and  
 CC thus reduce eosinophil numbers. The allergic diseases may be treated  
 CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,  
 CC it is possible that they may be used in the treatment of cancer and  
 CC helminthic infections.

XX SQ Sequence 116 AA;

Query Match 50.0%; Score 81; DB 21; Length 116;  
 Best Local Similarity 76.2%; Pred. No. 0.00045;  
 Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 SSGPSLQYIKANSKFIGITEL 31  
 :| :|||||  
 Db 24 TSNETMQYIKANSKFIGITEL 44

Search completed: September 24, 2003, 17:40:06  
 Job time : 35.036 secs



PS Example 4; Page -; 220pp; English.

PT Modified human tumour necrosis factor-alpha - comprises

CC The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunogenic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancers), endometriosis, uterine fibroids, benign prostatic  
 CC hyperplasia and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.

XX Sequence 50 AA;  
 SQ Query Match 53.7%; Score 87; DB 23; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-05; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSGPSL 16  
 Db 2 HWSYGLRPGSGPSL 16  
 |||||

RESULT 9  
 AAB20147  
 ID AAB20147 standard; Protein; 109 AA.  
 XX AC AAB20147;  
 XX DT 30-APR-2001 (first entry)  
 XX DE Growth differentiation factor 8 AutoVac construct GDF-8 P2-3.  
 XX KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;  
 XX KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;  
 XX KW cardiant; human; mutant; mutein.  
 XX OS Chimeric - Homo sapiens.  
 XX OS Chimeric - Clostridium tetani.  
 XX OS Synthetic.

XX FH Key Location/Qualifiers  
 XX FT Region 1..82 /note= "identical to residues 267-348 of human  
 XX FT GDF-8"  
 XX FT Region 83..97 /note= "tetanus toxoid P2 epitope"  
 XX FT Region 98..109 /note= "identical to residues 364-375 of human  
 XX FT GDF-8"  
 XX FT Misc-difference 73 /note= "Cys-73 may be substituted by Ser to avoid  
 XX FT disulfide bond formation"  
 XX FT Misc-difference 90..91 /note= "optionally replaced by Glu-Gly"  
 XX FT  
 XX PN WO200105820-A2.  
 XX XX 25-JAN-2001.  
 XX PD 20-JUL-2000; 2000WO-DK00413.  
 XX PF 20-JUL-1999; 99DK-0001014.  
 XX PR 26-JUL-1999; 99US-0145275.  
 XX XX (MEBI-) M & E BIOTECH AS.  
 XX PA Halkier T, Mouritsen S, Klysner S;  
 XX PI WPI; 2001-112680/12.  
 XX DR Increasing the muscle mass of animals used in meat production by down  
 XX PT

PT regulating growth differentiation factor 8 (GDF-8) activity in the  
 PT animal through induction of anti-GDF-8 antibody production -  
 XX Example 1; Page 99; 110pp; English.  
 PS  
 XX The present sequence is that of AutoVac construct GDF-8 P2-3,  
 CC comprising the 109 C-terminal amino acid residues of human  
 CC growth differentiation factor 8 (GDF-8) in which residues 83-97 are  
 CC replaced by the promiscuous tetanus toxin T-cell epitope P2 (see  
 CC AAB20143). It is an object of the invention to produce a  
 CC recombinant therapeutic vaccine that is capable of effecting  
 CC down-regulation of GDF-8 in order to increase the muscle growth  
 CC rate of farm animals. The vaccines (see AAB20143-53) are capable  
 CC of breaking autotolerance against autologous GDF-8. They comprise  
 CC the C-terminal portion of human GDF-8 in which a portion of the  
 CC native sequence is replaced by a T-cell epitope such as P2, with  
 CC minimal disturbance of the authentic 3-dimensional structure of  
 CC the protein. Nucleic acids encoding the GDF-8 variants can be used  
 CC for genetic immunisation of the animals. Down-regulation of GDF-8  
 CC activity can increase muscle mass by up to at least 45% in cattle,  
 CC pigs and poultry used for meat production, reducing the need for  
 CC antibiotic feed-additives. Anti-GDF8 vaccines can be used to  
 CC treat human diseases such as cancer cachexia where muscle atrophy is  
 CC pronounced and for patients suffering from acute and chronic heart  
 CC failure.

XX SQ Sequence 109 AA;  
 Query Match 52.5%; Score 85; DB 22; Length 109;  
 Best Local Similarity 63.3%; Pred. No. 0.00011;  
 Matches 19; Conservative 2; Mismatches 1; Indels 8; Gaps 1;

Qy 10 GSSGP-----SLQYKANSKFIGITEL 31  
 Db 68 GSAGPCTPTKMSPIQYKANSKFIGITEL 97  
 ||||| :|||||

RESULT 10  
 AAB20147  
 ID AAB20147 standard; Peptide; 24 AA.  
 XX AC AAB20147;  
 XX DT 11-SEP-2002 (first entry)  
 XX DE HBV antigen associated peptide #8.  
 XX KW Hepatitis B virus; HBV; antigen; major histocompatibility complex; MHC;  
 XX KW cytotoxic T cell; helper T cell; virucide; hepatotropic; immunogenic;  
 XX KW cytotoxic T lymphocyte; CTL; HLA-restricted response.  
 XX OS Synthetic.  
 XX PN US6322789-B1.  
 XX PD 27-NOV-2001.  
 XX XX 05-JUN-1995; 95US-0464496.  
 XX XX 26-AUG-1992; 92US-0935811.  
 XX PR 26-AUG-1991; 91US-0749568.  
 XX PR 29-JAN-1992; 92US-0827682.  
 XX PR 27-APR-1992; 92US-0874491.  
 XX XX (EPIM-) EPIMMUNE INC.  
 XX PI Vitello MA, Chesnut RW;  
 XX XX WPI; 2002-497942/53.  
 XX DR Immunogenic compositions for protecting against hepatitis B virus  
 XX PT infection -  
 XX PT

Query Match 54.9%; Score 89; DB 23; Length 51;  
 Best Local Similarity 63.3%; Pred. No. 1.2e-05;  
 Matches 19; Conservative 2; Mismatches 3; Indels 6; Gaps 1;

QY 2 HWSYGLRPGSSGPSL-----QYIKANSKF 25  
 ||||| ||||| : ||: |  
 Db 2 HWSYGLRPGSSGPSLDEKRIAKWEKASSVF 31

RESULT 7  
 AAU11425  
 ID AAU11425 standard; peptide; 37 AA.  
 XX  
 AC AAU11425;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Synthetic immunogen peptide 6.  
 XX  
 KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Clostridium tetani.  
 OS Chimeric - Mammalia.  
 OS Synthetic.

Key Location/Qualifiers  
 Misc-difference 1  
 /label= OTHER  
 /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
 Peptide 1..10  
 /note= "Gonadotrophin releasing hormone epitope"  
 Peptide 11..16  
 /note= "Spacer peptide"  
 Peptide 17..37  
 /note= "Tetanus toxoid sequence (947-967 aa)"  
 WO200185763-A2.  
 15-NOV-2001.  
 04-MAY-2001; 2001WO-US14363.  
 05-MAY-2000; 2000US-202328P.  
 (APHT-) APHTON CORP.  
 Grimes S, Michaeli D, Stevens VC;  
 WPI; 2002-049440/06.  
 Novel synthetic immunogen for inducing immune response against  
 gonadotrophin releasing hormone, comprises fusion peptide having  
 promiscuous helper T-cell peptide epitope and immunomimic peptide  
 epitope or its analogue  
 Claim 11; Page 9; 43pp; English.

The invention relates to a synthetic immunogen for inducing specific  
 antibodies against gonadotrophin releasing hormone (GnRH) also known  
 as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 peptide which comprises a promiscuous helper T-cell peptide epitope and  
 immunomimic peptide epitope or its analogue. The synthetic  
 immunogen is useful inducing an immune response against GnRH in an  
 animal subject, and as such is useful as a contraceptive and in the  
 treatment of diseases such as cancer (of the breast, uterus and other  
 gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 high and specific anti-GnRH antibody titres. The present sequence

CC is a synthetic immunogen of the invention.  
 XX  
 SQ Sequence 37 AA;  
 Query Match 53.7%; Score 87; DB 23; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSGPSL 16  
 ||||| ||||| |||||  
 Db 2 HWSYGLRPGSSGPSL 16

RESULT 8  
 AAU11429  
 ID AAU11429 standard; peptide; 50 AA.  
 XX  
 AC AAU11429;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Synthetic immunogen peptide 10.  
 XX  
 KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Clostridium tetani.  
 OS Chimeric - Mammalia.  
 OS Synthetic.

Key Location/Qualifiers  
 Misc-difference 1  
 /label= OTHER  
 /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
 Peptide 1..10  
 /note= "Gonadotrophin releasing hormone epitope  
 (1..10 aa)"  
 Peptide 11..16  
 /note= "Spacer peptide"  
 Peptide 17..37  
 /note= "Tetanus toxoid (947-967 aa)"  
 Peptide 38..41  
 /note= "Spacer peptide"  
 Peptide 42..50  
 /note= "Gonadotrophin releasing hormone epitope  
 (2-10 aa)"  
 Modified-site 50  
 /note= "Amidated glycine or glycine amide"  
 WO200185763-A2.  
 15-NOV-2001.  
 04-MAY-2001; 2001WO-US14363.  
 05-MAY-2000; 2000US-202328P.  
 (APHT-) APHTON CORP.  
 Grimes S, Michaeli D, Stevens VC;  
 WPI; 2002-049440/06.  
 Novel synthetic immunogen for inducing immune response against  
 gonadotrophin releasing hormone, comprises fusion peptide having  
 promiscuous helper T-cell peptide epitope and immunomimic peptide  
 epitope or its analogue  
 Claim 11; Page 11; 43pp; English.

KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Plasmodium falciparum.  
 OS Chimeric - Mammalia.  
 OS Synthetic.

# Key Location/Qualifiers

FT Misc-difference 1  
 FT /label= OTHER  
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
 FT Peptide  
 FT 1..10  
 FT /note= "Gonadotrophin releasing hormone epitope  
 FT (1..10 aa)"  
 FT Peptide  
 FT 11..16  
 FT /note= "Spacer peptide"  
 FT Peptide  
 FT 17..34  
 FT /note= "Malaria CSP protein (288-302 aa)"  
 FT Peptide  
 FT 35..38  
 FT /note= "Spacer peptide"  
 FT Peptide  
 FT 39..47  
 FT /note= "Gonadotrophin releasing hormone epitope  
 FT (2-10 aa)"  
 FT Modified-site  
 FT 47  
 FT /note= "Amidated glycine or glycineamide"

XX WO200185763-A2.

XX 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US14363.

XX 05-MAY-2000; 2000US-202328P.

XX (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;

XX WPT; 2002-049440/06.

XX Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue

XX Claim 11; Page 11; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.

XX Sequence 47 AA;

Query Match 54.9%; Score 89; DB 23; Length 47;

Best Local Similarity 83.3%; Pred. No. 1.1e-05;

Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSGPSLQYI 19

Db 2 HWSYGLRPGSGPSLKL 19

RESULT 6

AAU11431

ID AAU11431 standard; peptide; 51 AA.

XX AC AAU11431;

XX DT 12-MAR-2002 (first entry)

XX DE Synthetic immunogen peptide 12.

XX KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

OS Chimeric - Plasmodium falciparum.

OS Chimeric - Mammalia.  
 OS Synthetic.

# Key Location/Qualifiers

FT Misc-difference 1  
 FT /label= OTHER  
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
 FT Peptide  
 FT 1..10  
 FT /note= "Gonadotrophin releasing hormone epitope  
 FT (1..10 aa)"  
 FT Peptide  
 FT 11..16  
 FT /note= "Spacer peptide"  
 FT Peptide  
 FT 17..36  
 FT /note= "Malaria CSP protein (378-398 aa)"  
 FT Peptide  
 FT 37..42  
 FT /note= "Spacer peptide"  
 FT Peptide  
 FT 43..51  
 FT /note= "Gonadotrophin releasing hormone epitope  
 FT (2-10 aa)"  
 FT Modified-site  
 FT 51  
 FT /note= "Amidated glycine or glycineamide"

XX WO200185763-A2.

XX 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US14363.

XX 05-MAY-2000; 2000US-202328P.

XX (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;

XX WPT; 2002-049440/06.

XX Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue

XX Claim 11; Page 12-13; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.

XX Sequence 51 AA;



```

XX 15-NOV-2001.
PD
XX
XX 04-MAY-2001; 2001WO-US14363.
PF
XX
XX 05-MAY-2000; 2000US-202328P.
PR
XX
XX (APHT-) APHTON CORP.
PA
XX
XX Grimes S, Michaeli D, Stevens VC;
PI
XX WPI; 2002-049440/06.
DR
XX
XX Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide
PT epitope or its analogue
XX
XX Claim 11; Page 10; 43pp; English.
PS
XX
XX The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH) also known
CC as luteinising hormone releasing hormone, LHRH) comprising a fusion
CC peptide which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic
CC immunogen is useful inducing an immune response against GnRH in an
CC animal subject, and as such is useful as a contraceptive and in the
CC treatment of diseases such as cancer (of the breast, uterus and other
CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic
CC hypertrophy and prostate cancer. The immunogen is effective in eliciting
CC high and specific anti-GnRH antibody titres. The present sequence
CC is a synthetic immunogen of the invention.
XX
XX Sequence 31 AA;
SQ
Query Match 99.4%; Score 161; DB 23; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRPGSGPSLQYIKANSKFIGITEL 31
Db 2 HWSYGLRPGSGPSLQYIKANSKFIGITEL 31
RESULT 2
AAU11430
ID AAU11430 standard; peptide: 46 AA.
AC
XX
XX AAU11430;
DT
XX 12-MAR-2002 (first entry)
DE
XX Synthetic immunogen peptide 11.
XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
XX luteinising hormone releasing hormone; LHRH; contraceptive;
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX Chimeric - Clostridium tetani.
OS
OS Chimeric - Mammalia.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 1
FT /label= OTHER
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT
FT Peptide 1..10
FT /note= "Gonadotropin releasing hormone epitope
FT (1..10 aa)"
FT
FT Peptide 11..16
FT /note= "Spacer peptide"
FT

```

```

FT Peptide 17..31
FT /note= "Tetanus toxoid (830-844 aa)"
FT Peptide 32..37
FT /note= "Spacer peptide"
FT Peptide 38..46
FT /note= "Gonadotropin releasing hormone epitope
FT (2-10 aa)"
FT Modified-site 46
FT /note= "Amidated glycine or glycylamide"
XX
XX WO200185763-A2.
PN
XX
XX 15-NOV-2001.
PD
XX
XX 04-MAY-2001; 2001WO-US14363.
PF
XX
XX 05-MAY-2000; 2000US-202328P.
PR
XX
XX (APHT-) APHTON CORP.
PA
XX
XX Grimes S, Michaeli D, Stevens VC;
PI
XX WPI; 2002-049440/06.
DR
XX
XX Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide
PT epitope or its analogue
XX
XX Claim 11; Page 12; 43pp; English.
PS
XX
XX The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH) also known
CC as luteinising hormone releasing hormone, LHRH) comprising a fusion
CC peptide which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic
CC immunogen is useful inducing an immune response against GnRH in an
CC animal subject, and as such is useful as a contraceptive and in the
CC treatment of diseases such as cancer (of the breast, uterus and other
CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic
CC hypertrophy and prostate cancer. The immunogen is effective in eliciting
CC high and specific anti-GnRH antibody titres. The present sequence
CC is a synthetic immunogen of the invention.
XX
XX Sequence 46 AA;
SQ
Query Match 99.4%; Score 161; DB 23; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.5e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRPGSGPSLQYIKANSKFIGITEL 31
Db 2 HWSYGLRPGSGPSLQYIKANSKFIGITEL 31
RESULT 3
AAU11424
ID AAU11424 standard; peptide: 34 AA.
AC
XX
XX AAU11424;
DT
XX
XX 12-MAR-2002 (first entry)
DE
XX Synthetic immunogen peptide 5.
XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
XX luteinising hormone releasing hormone; LHRH; contraceptive;
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX Chimeric - Measles virus.
OS
OS Chimeric - Mammalia.

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:58:36 ; Search time 35.036 seconds  
(without alignments)  
140.442 Million cell updates/sec

Title: US-09-848-834A-15  
Perfect score: 162  
Sequence: 1 XHWSYGLRPGSSGSPSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
A_Geneseq_19Jun03.*			
1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*		
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	161	99.4	31	AAU11426 Synthetic immunoge
2	161	99.4	46	23 AAU11430 Synthetic immunoge
3	89	54.9	34	23 AAU11424 Synthetic immunoge
4	89	54.9	36	23 AAU11427 Synthetic immunoge
5	89	54.9	47	23 AAU11428 Synthetic immunoge
6	89	54.9	51	23 AAU11431 Synthetic immunoge
7	87	53.7	37	23 AAU11425 Synthetic immunoge
8	87	53.7	50	23 AAU11429 Synthetic immunoge
9	85	52.5	109	22 AAB20147 Growth differentia

10	84	51.9	24	23	ABP51515	HBV antigen associ
11	83	51.2	216	21	AA92665	MUC-1 analogue con
12	82.5	50.9	158	19	AAW81331	TNF2-7, a TNF-alpha
13	82.5	50.9	158	23	ABB07277	Human TNF-alpha an
14	81	50.0	116	21	AAAB45502	Modified murine in
15	80.5	49.7	158	19	AAAB45526	Modified murine in
16	80.5	49.7	158	19	AAW81327	TNF2-1, a TNF-alpha
17	80.5	49.7	158	23	ABB07280	Human TNF-alpha an
18	79	48.8	118	21	AAAB45491	Modified human int
19	79	48.8	118	21	AAAB45518	Modified human int
20	78	48.1	136	22	AAAB46190	Tetanus toxoid epi
21	78	48.1	136	22	AAAB49089	Amyloid beta tetan
22	78	48.1	182	21	AAAY84424	An osteoprotegerin
23	77	47.5	173	21	AAAY84425	DNA encoding osteo
24	76	46.9	16	23	AAU11413	Tetanus toxoid pre
25	76	46.9	19	21	AAAY9055	HLA Class II bindi
26	76	46.9	37	15	AAAR65389	Universal immunost
27	76	46.9	37	15	AAAR65383	Universal immunost
28	76	46.9	47	15	AAAR62723	LHRH-containing im
29	76	46.9	124	21	AAAB45492	Modified human int
30	76	46.9	124	21	AAAB45517	Modified human int
31	76	46.9	160	22	AAAB20153	Growth differentia
32	76	46.9	573	8	AAAP70345	Portion of B fragm
33	76	46.9	1315	22	AAAB61169	Clostridium tetani
34	75	46.3	25	21	AAAY92650	PSMpep007 - p2 ins
35	75	46.3	693	21	AAAY92647	Mutant human PSM a
36	75	46.3	750	21	AAAY92628	Mutant human prost
37	75	46.3	750	21	AAAY92637	Mutant human prost
38	75	46.3	750	21	AAAY92644	Mutant human prost
39	74	45.7	15	11	AAAR06310	Tetanus toxin epit
40	74	45.7	15	18	AAW35506	Tetanus toxin uni
41	74	45.7	15	18	AAW11505	Tetanus toxin epi
42	74	45.7	15	19	AAW67033	Tetanus toxin frag
43	74	45.7	15	19	AAW71321	Universal helper T
44	74	45.7	15	20	AAAY04051	T-helper epitope f
45	74	45.7	15	20	AAW67578	T-cell epitope pep

ALIGNMENTS

RESULT 1	AAU11426	
ID	AAU11426 standard; peptide; 31 AA.	
XX		
AC	AAU11426;	
XX		
XX	12-MAR-2002 (first entry)	
DT		
DE	Synthetic immunogen peptide 7.	
XX		
KW	Gonadotropin releasing hormone; GnRH; synthetic immunogen;	
KW	lutensising hormone releasing hormone; LHRH; contraceptive;	
KW	promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;	
KW	breast cancer; uterine cancer; gynaecological cancer; endometriosis;	
KW	uterine fibroid; benign prostatic hypertrophy; prostate cancer.	
OS	Chimeric - Clostridium tetani.	
OS	Chimeric - Mammalia.	
OS	Synthetic.	
XX		
Key	Location/Qualifiers	
FT	Misc-difference 1	
FT	/label= OTHER	
FT	/note= "Other= Pyro-glutamic acid or 5-oxo proline"	
FT	1..10	
FT	/note= "Gonadotropin releasing hormone epitope"	
FT	11..16	
FT	/note= "Spacer peptide"	
FT	17..31	
FT	/note= "Tetanus toxoid sequence (830-844 aa)"	
XX		
PN	WO200185763-A2.	

FT MOD\_RES 32 32  
 FT AMIDATION (G-33 PROVIDE AMIDE GROUP)  
 FT (BY SIMILARITY).  
 SQ SEQUENCE 95 AA; 10411 MW; 980C6988FC279BFC CRC64;  
 Query Match 32.1%; Score 52; DB 1; Length 95;  
 Best Local Similarity 88.9%; Pred. No. 0.87;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 HWSYGLRPG 10  
 |||||  
 Db 24 HWSYGLSPG 32

Search completed: September 24, 2003, 17:41:20  
 Job time : 6.85346 secs

Query Match 33.3%; Score 54; DB 1; Length 92;  
 Best Local Similarity 88.9%; Pred. No. 0.42;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
 DB 25 HWSYGLQPG 33

RESULT 14  
 GONL\_HAPBU STANDARD; PRT; 94 AA.  
 AC P51918; 093387;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)  
 DE (LH-RH I) (Luliberin I).  
 GN GNRH1.  
 OS Haplochromis burtoni (Burton's mouthbrooder).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;  
 OC Cichlidae; Astotilapia.  
 OC NCBI\_TaxID=8153;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95396797; PubMed=7667296;  
 RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;  
 RT "Three gonadotropin-releasing hormone genes in one organism suggest novel roles for an ancient peptide."  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99061842; PubMed=9843638;  
 RA White R.B., Fernald R.D.;  
 RT "Ontogeny of gonadotropin-releasing hormone (GnRH) gene expression reveals a distinct origin for GnRH-containing neurons in the midbrain."  
 RL Gen. Comp. Endocrinol. 112:322-329(1998).  
 RN [3]  
 RP SEQUENCE OF 23-32.  
 RC TISSUE=Pituitary;  
 RX MEDLINE=95372591; PubMed=7644702;  
 RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,  
 RA White S.A., Francis R.C., Fernald R.D., Licht P., Warby C.,  
 RA Sherwood N.M.;  
 RT "Primary structure of solitary form of gonadotropin-releasing hormone (GnRH) in cichlid pituitary; three forms of GnRH in brain of cichlid and pumpkinseed fish."  
 RL Regul. Pept. 57:43-53(1995).  
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-GONADAL AXIS.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPHYSAL AXONS.  
 CC -!- MASS SPECTROMETRY: MW=1113.9; METHOD=MALDI; RANGE=23-32.  
 CC -!- SIMILARITY: Belongs to the GnRH family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U31865; AAC59691.1; -.  
 CC EMBL; AF076961; AAC27716.1; -.  
 CC FIR; I50739; I50739.  
 CC InterPro; IPR002012; GnRH.  
 CC InterPro; IPR004079; GonadoliberinI.

DR Pfam; PF00446; GNRH; 1.  
 DR PRINTS; PR01541; GONADOLIBERNI.  
 DR PROSITE; PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal; Multigene family; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 22  
 FT CHAIN 23 94  
 FT PEPTIDE 23 32  
 FT PEPTIDE 36 94  
 FT MOD\_RES 23 23  
 FT MOD\_RES 32 32  
 FT CONFLICT 86 94  
 SQ SEQUENCE 94 AA; 10382 MW; E57DBA8333278D7 CRC64;  
 Query Match 32.1%; Score 52; DB 1; Length 94;  
 Best Local Similarity 88.9%; Pred. No. 0.86;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
 DB 24 HWSYGLSPG 32

RESULT 15  
 GONL\_MORSA STANDARD; PRT; 95 AA.  
 AC O73812;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)  
 DE (LH-RH I) (Luliberin I).  
 GN GNRH1.  
 OS Morone saxatilis (Striped bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidel;  
 OC Moronidae; Morone.  
 OC NCBI\_TaxID=34816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99061809; PubMed=9845669;  
 RA Chow M.M., Kight K.E., Gothilf Y., Alok D., Stubblefield J., Zohar Y.;  
 RT "Multiple GnRHs present in a teleost species are encoded by separate genes: analysis of the sbGnRH and cGnRH-II genes from the striped bass, Morone saxatilis."  
 RL J. Mol. Endocrinol. 21:277-289(1998).  
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the GnRH family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF056314; AAD03817.1; -.  
 CC InterPro; IPR002012; GnRH.  
 CC InterPro; IPR004079; GonadoliberinI.  
 DR Pfam; PF00446; GNRH; 1.  
 DR PROSITE; PR01541; GONADOLIBERNI.  
 DR PROSITE; PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal; Multigene family; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 22  
 FT CHAIN 23 95  
 FT PEPTIDE 23 32  
 FT PEPTIDE 36 95  
 FT MOD\_RES 23 23



```

RL Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
   luteinizing hormone-releasing hormone mRNA.";
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
   DNA locus";
RL Science 235:1514-1517(1987).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC -!- THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S50870; AAB24572.1; -
DR EMBL; M12579; AAA41263.1; -
DR EMBL; M31670; AAA41264.1; -
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -
DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
DR PIR; A40147; RHRTG.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyroliidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;
Query Match 35.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33
RESULT 10
GONI_TUPGB
ID GONI_TUPGB STANDARD; PRT; 92 AA.
AC Q95335;
DT 15-DEC-1998 (Rel. 37; Created)
DT 15-DEC-1998 (Rel. 37; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)]

```

```

DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
GN GnRH1 OR GnRH.
OS Tupaia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxID=9396;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGnRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GnRH gene expression in a
RT placental mammal.";
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC -!- THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U63326; AAB16837.1; -
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyroliidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;
Query Match 35.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33
RESULT 11
VG84_BPML5
ID VG84_BPML5 STANDARD; PRT; 66 AA.
AC O05301;
DT 01-FEB-1994 (Rel. 28; Created)
DT 01-FEB-1994 (Rel. 28; Last sequence update)
DT 01-FEB-1994 (Rel. 28; Last annotation update)
DE Gene 84 protein (GP84).
OS 84.
OS Mycobacteriophage L5.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC L5-like viruses.
OX NCBI_TaxID=31757;

```

DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Progonadoliblerin I precursor [Contains: Gonadoliblerin I (LH-RH I)  
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
DE hormone I) (GNRH I) (Luliberin I) (Gonadorelin); GNRH-associated  
DE peptide I].  
DN GNRH1 OR GNRH OR LHRH.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89366682; PubMed=2671939;  
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;  
RT "The complete nucleotide sequence of the human gonadotropin-releasing  
RT hormone gene.";  
RL Nucleic Acids Res. 17:6403-6403(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86094338; PubMed=2867548;  
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;  
RT "Isolation of the gene and hypothalamic cDNA for the common precursor  
RT of gonadotropin-releasing hormone and prolactin release-inhibiting  
RT factor in human and rat.";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANT SER-16.  
RX MEDLINE=85012739; PubMed=6090951;  
RA Seeburg P.H., Adelman J.P.;  
RT "Characterization of cDNA for precursor of human luteinizing hormone  
RT releasing hormone.";  
RL Nature 311:666-668(1984).  
RN [4]  
RP SEQUENCE OF 24-33.  
RX MEDLINE=83126573; PubMed=6760865;  
RA Tan L., Rousseau P.;  
RT "The chemical identity of the immunoreactive LHRH-like peptide  
RT biosynthesized in the human placenta.";  
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).  
RN [5]  
RP VARIANT SER-16.  
RX MEDLINE=99318093; PubMed=10391209;  
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,  
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
RA Lander E.S.;  
RT "Characterization of single-nucleotide polymorphisms in coding regions  
RT of human genes.";  
RL Nat. Genet. 22:231-238(1999).  
RN [6]  
RP ERRATUM.  
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,  
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
RA Lander E.S.;  
RL Nat. Genet. 23:373-373(1999).  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),  
CC Lutrepulse or Lutrelief (Ferring Pharmaceuticals) and Relisorm  
CC (Serono).  
CC -1- SIMILARITY: Belongs to the GNRH family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL; X01059; CAA25526.1; -;  
DR EMBL; M12578; AAA35916.1; -;  
DR EMBL; X15215; CAA33285.1; -;  
DR PIR; S05308; RHHUG.  
DR Gene; HGNC:4419; GNRH1.  
DR MIN; 152760; -;  
DR GO; GO:0005625; C:soluble fraction; TAS.  
DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.  
DR GO; GO:0007267; P:cell-cell signaling; TAS.  
DR GO; GO:0007275; P:development; TAS.  
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
DR GO; GO:0007165; P:signal transduction; TAS.  
DR InterPro; IPR002012; GNRH.  
DR InterPro; IPR004079; GonadoliblerinI.  
DR Pfam; PF00446; GNRH; 1.  
DR PRINTS; PR01541; GONADOLIBRNI.  
DR PROSITE; PS00473; GNRH; 1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Placenta; Pharmaceutical; Signal; Polymorphism;  
KW Pyrolidone carboxylic acid.  
FT SIGNAL 1 23 PROGNADOLIBERIN I.  
FT CHAIN 24 92 GONADOLIBERIN I.  
FT PEPTIDE 24 33 GNRH-ASSOCIATED PEPTIDE I.  
FT PEPTIDE 37 92 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
FT ACT\_SITE 26 26 ACTIVITY.  
FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
FT VARIANT 16 16 W -> S (IN DBSNP:6185).  
FT /FTID=VAR\_013943.  
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;  
Query Match 35.8%; Score 58; DB 1; Length 92;  
Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;  
Oy 2 HWSYGLRPG 10  
Db 25 HWSYGLRPG 33  
RESULT 9  
ID GONL\_RAT STANDARD; PRT; 92 AA.  
AC P07490;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Progonadoliblerin I precursor [Contains: Gonadoliblerin I (LH-RH I)  
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
DE hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor  
DE I].  
DN GNRH1 OR GNRH.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86094338; PubMed=2867548;  
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;  
RT "Isolation of the gene and hypothalamic cDNA for the common precursor  
RT of gonadotropin-releasing hormone and prolactin release-inhibiting  
RT factor in human and rat.";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89384661; PubMed=2476669;  
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;  
RT "The rat gonadotropin-releasing hormone: SH locus: structure and  
RT hypothalamic expression.";

DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor  
DE I].  
DE GN GNRH1 OR GNRH.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87069928; PubMed=3024317;  
RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,  
RA Phillips H.S., Nikolics K., Seeburg P.H.;  
RT "A deletion truncating the gonadotropin-releasing hormone gene is  
RT responsible for hypogonadism in the hpg mouse.";  
RL Science 234:1366-1371(1986).  
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the GNRH family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M14872; AAA37717.1; -.  
CC PIR; A47578; RHMSG.  
CC MGD; MGI:95789; GnRh.  
CC InterPro; IPR002012; GnRH.  
CC InterPro; IPR004079; GonadoliberinI.  
CC Pfam; PF00446; GnRH; 1.  
CC PRINTS; PR01541; GONADOLIBERNI.  
CC PROSITE; PS00473; GNRH; 1.  
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
CC Placenta; Signal; Pyrrolidone carboxylic acid.  
CC FT SIGNAL 1 21  
CC FT CHAIN 22 90 PROGNADOLIBERIN I.  
CC FT PEPTIDE 22 31 GONADOLIBERIN I.  
CC FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.  
CC FT ACT\_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
CC FT ACTIVITY.  
CC FT MOD\_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.  
CC FT MOD\_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).  
CC FT SEQUENCE 90 AA; 10337 MW; 1C0766FA826E4D9 CRC64;  
CC  
CC Query Match 35.8%; Score 58; DB 1; Length 90;  
CC Best Local Similarity 100.0%; Pred. No. 0.1;  
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CC  
CC QY 2 HWSYGLRPG 10  
CC Db |||||  
CC 23 HWSYGLRPG 31  
CC  
CC RESULT 7  
CC GONL\_PIG STANDARD; PRT; 91 AA.  
CC AC P49921;  
CC DT 01-OCT-1996 (Rel. 34, Created)  
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)  
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
CC DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
CC (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
CC hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide II.  
CC GNRH1 OR GNRH.  
CC Sus scrofa (Pig).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Hypothalamus;  
RA Weesner G.D., Matteri R.L., Becker B.A.;  
RN Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 24-33.  
RX MEDLINE=72114303; PubMed=4946067;  
RA Baba Y., Matsuo H., Schally A.V.;  
RT "Structure of the porcine LH- and FSH-releasing hormone. II.  
RT Confirmation of the proposed structure by conventional sequential  
RT analyses.";  
RL Biochem. Biophys. Res. Commun. 44:459-463(1971).  
RN [3]  
RP SYNTHESIS OF GONADOLIBERIN.  
RX MEDLINE=72065376; PubMed=4942726;  
RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;  
RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-  
RT phase method.";  
RL Biochem. Biophys. Res. Commun. 45:822-827(1971).  
RN [4]  
RP SYNTHESIS OF GONADOLIBERIN.  
RX MEDLINE=72117544; PubMed=4946275;  
RA Baba Y., Arimura A., Schally A.V.;  
RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";  
RL Biochem. Biophys. Res. Commun. 45:483-487(1971).  
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the GNRH family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L32864; AAA31066.1; -.  
CC InterPro; IPR002012; GnRH.  
CC InterPro; IPR004079; GonadoliberinI.  
CC Pfam; PF00446; GnRH; 1.  
CC PRINTS; PR01541; GONADOLIBERNI.  
CC PROSITE; PS00473; GNRH; 1.  
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
CC Placenta; Signal; Pyrrolidone carboxylic acid.  
CC FT SIGNAL 1 23  
CC FT CHAIN 24 91 PROGNADOLIBERIN I.  
CC FT PEPTIDE 24 33 GONADOLIBERIN I.  
CC FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.  
CC FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
CC FT ACTIVITY.  
CC FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
CC FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
CC FT SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;  
CC  
CC Query Match 35.8%; Score 58; DB 1; Length 91;  
CC Best Local Similarity 100.0%; Pred. No. 0.11;  
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CC  
CC QY 2 HWSYGLRPG 10  
CC Db |||||  
CC 25 HWSYGLRPG 33  
CC  
CC RESULT 8  
CC GONL\_HUMAN STANDARD; PRT; 92 AA.  
CC ID GONL\_HUMAN  
CC AC P01148;

Db 2 HWSYGLRPGG-----KRNAK 16

```
RESULT 4
GONI_MACMU STANDARD; PRT; 67 AA.
ID GONI_MACMU STANDARD; PRT; 67 AA.
AC P55247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberein I precursor (Contains: Gonadoliberein I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I]
DE (Fragment).
GN GNRHI OR GNRH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female rhesus
RT macaques.";
RL Neuroendocrinology 60:346-359(1994).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GNRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S75918; AAB33096.1; -
CC PIR; I78541; I78541.
CC InterPro; IPR002012; GNRH.
CC Pfam; PF004079; GonadolibereinI.
CC PRINTS; PR01541; GONADOLIBERNI.
CC PROSITE; PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Signal; Pyrrolidone carboxylic acid.
CC NON_TER 1 1
CC SIGNAL <1 5 BY SIMILARITY.
CC CHAIN 6 >67 PROGONADOLIBERIN I.
CC PEPTIDE 6 15 GONADOLIBERIN I.
CC ACT_SITE 19 >67 GNRH-ASSOCIATED PEPTIDE I.
CC MOD_RES 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC ACTIVITY (BY SIMILARITY).
CC MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
CC SIMILARITY).
CC MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
CC SIMILARITY).
CC NON_TER 67 67
CC SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
Query Match 35.8%; Score 58; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRPG 10
Db 7 HWSYGLRPG 15
-----
RESULT 5
GONI_XENLA STANDARD; PRT; 89 AA.
ID GONI_XENLA STANDARD; PRT; 89 AA.
AC P45636;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberein I precursor (Gonadotropin-releasing hormone I) (GNRH-I)
DE (LH-RH) (Luliberin I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Forebrain;
RX MEDLINE=94185563; PubMed=8137750;
RA Hayes W.P., Gray S., Battey J.F.;
RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GNRH-associated peptide, but brain onset is delayed until
RT metamorphosis.";
RL Endocrinology 134:1835-1844(1994).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GNRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L28040; AAA49728.1; -
CC PIR; I51423; I51423.
CC InterPro; IPR002012; GNRH.
CC InterPro; IPR004079; GonadolibereinI.
CC Pfam; PF00446; GNRH; 1.
CC PRINTS; PR01541; GONADOLIBERNI.
CC PROSITE; PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Signal; Pyrrolidone carboxylic acid.
CC SIGNAL 1 23
CC CHAIN 24 89 PROGONADOLIBERIN I.
CC PEPTIDE 24 33 GONADOLIBERIN I.
CC PEPTIDE 37 89 GONADOTROPIN-RELEASING HORMONE ASSOCIATED
CC PEPTIDE.
CC FT PEPTIDE 37 85 GNRH-ASSOCIATED PEPTIDE I (GAP).
CC MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
CC MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
CC SEQUENCE 89 AA; 10246 MW; 6FAF36FBAE0D4284 CRC64;
Query Match 35.8%; Score 58; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33
-----
RESULT 6
GONI_MOUSE STANDARD; PRT; 90 AA.
ID GONI_MOUSE STANDARD; PRT; 90 AA.
AC P13562;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```

Query Match 46.9%; Score 76; DB 1; Length 1314;  
 Best Local Similarity 93.8%; Pred. No. 0.0044; 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31  
 :|||||

Db 828 MOYIKANSKFIGITEL 843

RESULT 2

GONI_MESAU	STANDARD;	PRT;	63 AA.
ID GONI_MESAU			
AC O09163;			
DT 15-DEC-1998 (Rel. 37, Created)			
DT 15-DEC-1998 (Rel. 37, Last sequence update)			
DT 28-FEB-2003 (Rel. 41, Last annotation update)			
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)]			
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]			
DE (Fragment).			
DE GnRH I OR GnRH OR LHRH.			
OS Mesocricetus auratus (Golden hamster).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;			
OC Mesocricetus.			
OX NCBI_TaxID=10036;			
RN [1]			
RP SEQUENCE FROM N.A.			
RA Jansen H.T., Stevens P.J., Zeidler P., Lehman M.N.;			
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.			
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.			
CC -!- SUBCELLULAR LOCATION: Secreted.			
CC -!- SIMILARITY: Belongs to the GnRH family.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement/">http://www.isb-sib.ch/announcement/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC EMBL; U01938; AAB51302.1; -			
DR InterPro; IPR002012; GnRH.			
DR InterPro; IPR004079; GonadolibereinI.			
DR Pfam; PF00446; GnRH; 1.			
DR PRINTS; PR01541; GONADOLIBERNI.			
DR PROSITE; PS00473; GnRH; 1.			
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;			
KW Placenta; Pyrrolidone carboxylic acid.			
FT NON_TER 1 1			
FT CHAIN 1 >63			
FT PEPTIDE 1 10			
FT PEPTIDE 14 >63			
FT ACT_SITE 3 3			
FT MOD_RES 1 1			
FT MOD_RES 10 10			
FT MOD_RES 63 63			
FT NON_TER 63 63			
SQ SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;			

Query Match 37.7%; Score 61; DB 1; Length 63;  
 Best Local Similarity 46.2%; Pred. No. 0.025;  
 Matches 12; Conservative 3; Mismatches 3; Indels 8; Gaps 1;

QY 2 HWSYGLRPGSGPSLOYIKANSKFIG 27  
 :|||||

Db 2 HWSYGLRPGG-----KRNAERLG 19

RESULT 3

GONI_SHEEP	STANDARD;	PRT;	61 AA.
ID GONI_SHEEP			
AC Q28588;			
DT 15-DEC-1998 (Rel. 37, Created)			
DT 15-DEC-1998 (Rel. 37, Last sequence update)			
DT 28-FEB-2003 (Rel. 41, Last annotation update)			
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)]			
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]			
DE (Fragment).			
DE GnRH I OR GnRH OR LHRH.			
OS Ovis aries (Sheep).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Chordartiodactyla; Ruminantia; Pecora; Bovidae;			
OC Bovidae; Caprinae; Ovis.			
OX NCBI_TaxID=9940;			
RN [1]			
RP SEQUENCE OF 12-61 FROM N.A.			
RC STRAIN=Western range; TISSUE=Hypothalamus;			
RA Rodriguez R.E., Wise M.E.;			
RL Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.			
RN [2]			
RP SEQUENCE OF 1-10.			
RX MEDLINE=72094314; PubMed=4550508;			
RA Burgus R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J., Fellows R., Blackwell R., Vale W., Guillemin R.;			
RT "Primary structure of the ovine hypothalamic luteinizing hormone-releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass spectrometry-decapeptide-Edman degradation)";			
RT Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).			
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.			
CC -!- SUBCELLULAR LOCATION: Secreted.			
CC -!- SIMILARITY: Belongs to the GnRH family.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement/">http://www.isb-sib.ch/announcement/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC EMBL; U02517; AAA03433.1; -			
DR InterPro; IPR002012; GnRH.			
DR InterPro; IPR004079; GonadolibereinI.			
DR Pfam; PF00446; GnRH; 1.			
DR PRINTS; PR01541; GONADOLIBERNI.			
DR PROSITE; PS00473; GnRH; 1.			
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;			
KW Placenta; Pyrrolidone carboxylic acid.			
FT NON_TER 1 1			
FT CHAIN 1 >61			
FT PEPTIDE 1 10			
FT PEPTIDE 14 >61			
FT ACT_SITE 3 3			
FT MOD_RES 1 1			
FT MOD_RES 10 10			
FT MOD_RES 61 61			
FT NON_TER 61 61			
SQ SEQUENCE 61 AA; 6828 MW; 63962AAE319B8F0 CRC64;			

Query Match 37.0%; Score 60; DB 1; Length 61;  
 Best Local Similarity 52.2%; Pred. No. 0.034;  
 Matches 12; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

QY 2 HWSYGLRPGSGPSLOYIKANSK 24  
 :|||||

RL Eur. J. Biochem. 188:39-45(1990).  
 RN [6]  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE-92037649; PubMed-1935979;  
 RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;  
 RT "Limited proteolysis of tetanus toxin. Relation to activity and  
 identification of cleavage sites.";  
 RL Eur. J. Biochem. 202:41-51(1991).  
 RN [7]  
 RP IDENTIFICATION AS ZINC-PROTEASE.  
 RX MEDLINE-93010948; PubMed-1396558;  
 RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,  
 RA Montecucco C.;  
 RT "Tetanus toxin is a zinc protein and its inhibition of  
 neurotransmitter release and protease activity depend on zinc.";  
 RL EMBO J. 11:3577-3583(1992).  
 RN [8]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE-93063293; PubMed-1331807;  
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,  
 RA Dasgupta B.R., Montecucco C.;  
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release  
 by proteolytic cleavage of synaptobrevin.";  
 RL Nature 359:832-835(1992).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.  
 RX MEDLINE-97475217; PubMed-9334741;  
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,  
 RA Sax M.;  
 RT "Structure of the receptor binding fragment HC of tetanus  
 neurotoxin.";  
 RL Nat. Struct. Biol. 4:788-792(1997).  
 CC -|- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77  
 BOND OF SYNAPTOSOMAL-2.  
 CC -|- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in  
 synaptobrevin 2.  
 CC -|- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -|- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO  
 YIELD SUCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE  
 AND ARE NON-TOXIC AFTER SEPARATION.  
 CC -|- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO  
 GANGLIOSIDE RECEPTORS.  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X04436; CAA28033.1; -;  
 DR EMBL; X06214; CAA29564.1; -;  
 DR EMBL; AF528097; AAO37454.1; -;  
 DR EMBL; M12739; AAA23282.1; -;  
 DR PIR; A25689; BTCLTN.  
 DR PDB; 1AF9; 29-APR-98.  
 DR PDB; 1A8D; 14-OCT-98.  
 DR PDB; 1D0H; 27-MAR-00.  
 DR PDB; 1DFO; 24-MAR-00.  
 DR PDB; 1DIW; 24-MAR-00.  
 DR PDB; 1DLL; 24-MAR-00.  
 DR PDB; 1FV3; 05-SEP-01.  
 DR MEROPS; M27.001; -;  
 DR InterPro; IPR000395; Bontoxilysin.  
 DR InterPro; IPR006025; Zn\_MTPeptidse.  
 DR Pfam; PF01742; Peptidase\_M27; 1.

DR PRINTS; PR00760; BONTOXILYSIN.  
 DR PRODOM; PD001963; Bontoxilysin; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;  
 FT 3D-structure; Complete proteome.  
 FT INIT\_MET 0 0  
 FT CHAIN 457 1 456 TETANUS TOXIN LIGHT CHAIN.  
 FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.  
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 233 233 BY SIMILARITY.  
 FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT TRANSMEM 226 246 POTENTIAL.  
 FT TRANSMEM 669 689 POTENTIAL.  
 FT DISULFID 438 466 INTERCHAIN.  
 FT DISULFID 1076 1092  
 FT HELIX 876 882  
 FT TURN 883 883  
 FT STRAND 884 891  
 FT TURN 892 893  
 FT STRAND 894 897  
 FT STRAND 904 907  
 FT TURN 909 910  
 FT STRAND 912 915  
 FT STRAND 920 925  
 FT TURN 928 929  
 FT STRAND 932 935  
 FT STRAND 938 940  
 FT HELIX 941 946  
 FT TURN 949 956  
 FT STRAND 962 968  
 FT TURN 969 970  
 FT STRAND 972 977  
 FT STRAND 980 981  
 FT HELIX 983 985  
 FT STRAND 987 995  
 FT TURN 996 997  
 FT STRAND 998 1004  
 FT TURN 1006 1007  
 FT STRAND 1010 1016  
 FT STRAND 1020 1020  
 FT TURN 1021 1022  
 FT STRAND 1031 1037  
 FT TURN 1039 1040  
 FT STRAND 1042 1047  
 FT TURN 1048 1049  
 FT STRAND 1050 1056  
 FT TURN 1058 1059  
 FT STRAND 1068 1074  
 FT TURN 1079 1080  
 FT STRAND 1082 1091  
 FT HELIX 1097 1105  
 FT TURN 1106 1107  
 FT STRAND 1112 1112  
 FT STRAND 1114 1114  
 FT TURN 1116 1117  
 FT STRAND 1120 1120  
 FT STRAND 1122 1122  
 FT TURN 1123 1124  
 FT STRAND 1127 1131  
 FT HELIX 1132 1134  
 FT TURN 1135 1136  
 FT STRAND 1137 1141  
 FT TURN 1144 1145  
 FT STRAND 1148 1152  
 FT STRAND 1155 1158  
 FT TURN 1159 1162  
 FT STRAND 1163 1166  
 FT STRAND 1173 1178  
 FT TURN 1184 1185  
 FT STRAND 1188 1188  
 FT STRAND 1190 1190  
 FT TURN 1191 1192  
 FT STRAND 1193 1201

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:00:31 : Search time 5.75346 Seconds  
(without alignments)  
253.383 Million cell updates/sec

Title: US-09-848-834A-15  
Perfect score: 162  
Sequence: 1 XHWSYGLRPGSSGSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	76	46.9	1314	1	TETX_CLOTE
2	61	37.7	63	1	GON1_MESAU
3	60	37.0	61	1	GON1_SHEEP
4	58	35.8	67	1	GON1_MACMU
5	58	35.8	89	1	GON1_XENLA
6	58	35.8	90	1	GON1_MOUSE
7	58	35.8	91	1	GON1_PIG
8	58	35.8	92	1	GON1_HUMAN
9	58	35.8	92	1	GON1_TUPGB
10	58	35.8	92	1	GON1_RAT
11	55.5	34.3	66	1	VG84_BPML5
12	54	33.3	10	1	GON1_ALLMI
13	54	33.3	92	1	GON1_CHICK
14	52	32.1	94	1	GON1_HAPBU
15	52	32.1	95	1	GON1_MORSA
16	52	32.1	95	1	GON1_PAGMA
17	52	32.1	95	1	GON1_SPAU
18	52	32.1	99	1	GON1_DICLA
19	52	32.1	575	1	ACBA_LYCES
20	50.5	31.2	90	1	GON3_DICLA
21	50	30.9	80	1	GON1_CLAGA
22	50	30.9	90	1	GON8_RANDY
23	50	30.9	249	1	PRA_MYCLE
24	49	30.2	92	1	GON1_CAVPO
25	49	30.2	190	1	HIS5_PYRAE
26	49	30.2	293	1	KHSE_PYRHO
27	49	30.2	294	1	KHSE_PYRAB
28	49	30.2	408	1	SEPR_THESR
29	49	30.2	444	1	GARP_ECOLI
30	49	30.2	485	1	RT16_MYXXA
31	48.5	29.9	892	1	IF2_YERPE
32	48	29.6	390	1	YB12_SCHPO
33	47.5	29.3	89	1	GON3_PORNO

34	47.5	29.3	90	1	GON3_HAPBU	P45652 haplochromi
35	47.5	29.3	90	1	GON3_PAGMA	P51921 pagrus majo
36	47.5	29.3	90	1	GON3_SPAU	P51923 sparus aura
37	47.5	29.3	276	1	SFAS_CHLRE	Q39618 chlamydomon
38	47.5	29.3	407	1	VG10_HSVSA	P24913 herpesvirus
39	47.5	29.3	526	1	MAUQ_CHLPN	Q92812 chlamydia p
40	47	29.0	10	1	GON1_CLUPA	P81749 clupea pall
41	47	29.0	120	1	HV03_MOUSE	P01747 mus musculu
42	47	29.0	140	1	HV02_MOUSE	P01746 mus musculu
43	47	29.0	410	1	RLMS_EMENI	P02382 emericeia
44	47	29.0	494	1	ILVC_VIBPA	Q87kn4 vibrio para
45	47	29.0	494	1	ILVC_VIBVU	Q8ddc8 vibrio vuln

ALIGNMENTS

RESULT 1

TETX\_CLOTE

ID TETX\_CLOTE STANDARD; PRT; 1314 AA.

AC P04958;

DT 13-AUG-1987 (Rel. 05, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Tetanus toxin precursor (EC 3.4.24.68) (tentoxylisin) [Contains:

DE Tetanus toxin light chain (Tetanus toxin chain D); Tetanus toxin heavy

DE chain (Tetanus toxin chain H)].

GN TETX OR CTP60.

OS Clostridium tetani.

OG Plasmid pE88, and Plasmid 75 Kbp.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI\_Taxid=1513;

RN [1]

RP SEQUENCE FROM N.A.

RC PLASMID=75 Kbp;

RX MEDLINE=87053814; PubMed=3536478;

RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,

RA Weller U., Hudei M., Habermann E., Niemann H.;

RT "Tetanus toxin: primary structure, expression in E. coli, and

RT homology with botulinum toxins."

EMBO J. 5:2495-2502(1986).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CN3911; PLASMID=75 Kbp;

RX MEDLINE=87040747; PubMed=3774547;

RA Fairweather N.F., Lyness V.A.;

RT "The complete nucleotide sequence of tetanus toxin.";

Nucleic Acids Res. 14:7809-7812(1986).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Massachusetts / E88; PLASMID=pE88;

RX MEDLINE=22457253; PubMed=12552129;

RA Brueggemann H., Baeumer S., Fricke W.F., Wierse A., Liesegang H.,

RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,

RA Gottschalk G.;

RT "The genome sequence of Clostridium tetani, the causative agent of

RT tetanus disease.";

Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).

RN [4]

RP SEQUENCE OF 742-1314 FROM N.A.

RC PLASMID=75 Kbp;

RX MEDLINE=86085672; PubMed=3510187;

RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;

RT "Cloning, nucleotide sequencing, and expression of tetanus toxin

RT fragment C in Escherichia coli.";

J. Bacteriol. 165:21-27(1986).

RN [5]

RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.

RX MEDLINE=90201034; PubMed=2108021;

RA Krieglstein K., Henschen A., Weller U., Habermann E.;

RT "Arrangement of disulfide bridges and positions of sulphhydryl groups

RT in tetanus toxin.";



Db 2 HWSYGLQPG 10

# RESULT 13

I50644

gonadoliberin I precursor - chicken

N:Alternate names: gonadotropin-releasing hormone I

C:Species: Gallus gallus (chicken)

C>Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 16-Jul-1999

C:Accession: I50644; S33507

R:Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.

J. Mol. Endocrinol. 11, 19-29, 1993

A:Title: Characterization of the chicken preprogonadotropin-releasing hormone-I gene.

A:Reference number: I50644; MUID:94059355; PMID:7902095

A:Accession: I50644

A>Status: translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-92 <DU2>

A:Cross-references: EMBL:X69491; NID:g496326; PIDN:CAA49246.1; PID:g311612

C:Genetics:

A:Introns: 47/3; 79/3

C:Superfamily: gonadoliberin

Query Match 33.3%; Score 54; DB 2; Length 92;

Best Local Similarity 88.9%; Pred. NO. 1.2;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

|||||:|

Db 25 HWSYGLQPG 33

# RESULT 14

E95361

probable muconate cycloisomerase (EC 5.5.1.1) [imported] - Sinorhizobium meliloti (strain

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C:Accession: E95361

R:Barnett, M.J.; Jones, R.F.; Fisher, T.J.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: E95361

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <KUP>

A:Cross-references: GB:AE006469; PIDN:AAK65455.1; PID:g14523923; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: Sma1461

A:Genome: plasmid

C:Keywords: intramolecular lyase; isomerase

Query Match

33.0%; Score 53.5; DB 2; Length 374;

Best Local Similarity 44.8%; Pred. No. 7;

Matches 13; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

QY 2 HWSYGLRPGSGSLQVYKANSKFGITE 30

|||||:| | | | | | | | | | | | |

Db 21 HWSYGIRE-SFVAVNLIEADDTVGIGE 48

# RESULT 15

I50739

gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)

C:Species: Haplochromis burtoni

C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: I50739

R:White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.

Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995

A:Title: Three gonadotropin-releasing hormone genes in one organism suggest novel

A:Reference number: I50739; MUID:95396797; PMID:7667296

A:Accession: I50739

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-98 <WHI>

A:Cross-references: EMBL:U31865; NID:g905398; PIDN:AAC59691.1; PID:g905399

C:Superfamily: gonadoliberin

Query Match

32.1%; Score 52; DB 2; Length 98;

Best Local Similarity 88.9%; Pred. NO. 2.6;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

|||||:|

Db 24 HWSYGLSPG 32

Search completed: September 24, 2003, 17:49:34

Job time: 11.9058 secs

A;Residues: 1-92 <ADE>  
A;Cross-references: GB:M12579; NID:g204445; PIDN:AAA41263.1; PID:g204446  
R;Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Bialock, J.E.  
Cell. Mol. Neurobiol. 12, 447-454, 1992  
A;Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone  
A;Reference number: A48410; MUID:93105480; PMID:1468115  
A;Accession: A48410  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-92 <MAI>  
A;Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060  
A;Experimental source: thymus  
A;Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P:121083)  
C;Genetics:  
A;Introns: 47/3; 79/3  
C;Function:  
A;Description: stimulates pituitary secretion of lutropin and follitropin  
A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C;Superfamily: gonadoliberin  
C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid; 1  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;1-23/2/Product: progadoliberin #status predicted <PGN>  
F;24-33/Product: gonadoliberin #status predicted <GLN>  
F;37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>  
F;24/Modified site: pyroglutamic acid (Gln) (in mature form) #status predicted  
F;33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 35.8%; Score 58; DB 1; Length 92;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
|||||||  
DB 25 HWSYGLRPG 33

RESULT 9  
T52510  
hypothetical protein B2J23.60 [imported] - Neurospora crassa  
C;Species: Neurospora crassa  
C;Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000  
C;Accession: T52510  
R;Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, September 2000  
A;Reference number: T26053  
A;Accession: T52510  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-719 <SCH>  
A;Cross-references: EMBL:AL442164; GSPDB:GN00116; NCSP:B2J23.60  
A;Experimental source: BAC clone B2J23; strain OR74A  
C;Genetics:  
A;Gene: NCSP:B2J23.60  
A;Map position: 6  
A;Introns: 349/1; 601/1

Query Match 34.6%; Score 56; DB 2; Length 719;  
Best Local Similarity 45.0%; Pred. No. 6.3;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 WSYGLRPGSGPSLQYIKAN 22  
||||| : : : :  
DB 557 WSYSGRPGSAGGLMSFVSAS 576

RESULT 10  
S31029  
gene 84 protein - Mycobacterium phage L5  
C;Species: Mycobacterium phage L5  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 08-Oct-1999  
C;Accession: S31029  
R;Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.  
Mol. Microbiol. 7, 407-417, 1993

A;Title: Superinfection immunity of mycobacteriophage L5: applications for genetic tr  
A;Reference number: S30949; MUID:93211283; PMID:8459767  
A;Accession: S31029  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-66 <DON>  
A;Cross-references: EMBL:Z18946; NID:g15859; PIDN:CAA79460.1; PID:es9702; PID:g579152  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993;  
C;Genetics:  
A;Gene: 84  
A;Start codon: GTG

Query Match 34.3%; Score 55.5; DB 2; Length 66;  
Best Local Similarity 43.8%; Pred. No. 0.51;  
Matches 14; Conservative 2; Mismatches 7; Indels 9; Gaps 2;

QY 5 YGL-----RPGSSGPSLQYIKANSKFIGITEL 31  
||||| : : : :  
DB 36 YGFEVDWYEPGESG----YIKRNGKFGVTWEV 63

RESULT 11  
G48677  
Ig heavy chain V-D-J region (419.1) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 17-Mar-1999  
C;Accession: G48677  
R;Tassignon, J.; Brait, M.; Jamila, I.; Urbain, J.; Gottlieb, P.; Brown, A.; Haseman  
Proc. Natl. Acad. Sci. U.S.A. 90, 9508-9512, 1993  
A;Title: Molecular characterization of monoclonal CRI-A-positive anti-arsonate antibo  
A;Reference number: A48677; MUID:94022404; PMID:8415731  
A;Accession: G48677  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-123 <TAS>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 34.0%; Score 55; DB 2; Length 123;  
Best Local Similarity 40.0%; Pred. No. 1.2;  
Matches 16; Conservative 1; Mismatches 11; Indels 12; Gaps 2;

QY 4 SYGL-----RPGSS-----GPSLQYIKANSKFIGITEL 31  
||||| : : : :  
DB 31 SYGVNVWKQRPQGGLEWIGYINFGNDYIKYNEKFKGTTL 70

RESULT 12  
RHA01  
gonadoliberin I - American alligator  
N;Alternate names: gonadotropin-releasing hormone I  
C;Species: Alligator mississippiensis (American alligator)  
C;Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Mar-1997  
C;Accession: A60066  
R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan  
Regul. Pept. 33, 105-116, 1991  
A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains  
A;Reference number: A60066; MUID:91352338; PMID:1882082  
A;Accession: A60066  
A;Molecule type: protein  
A;Residues: 1-10 <LOW>  
C;Superfamily: gonadoliberin  
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F;1/Modified site: pyroglutamic acid (Gln) #status experimental  
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 33.3%; Score 54; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.1;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
|||||||

A:Gene: GnRH-I  
C:Superfamily: gonadoliberin

Query Match 35.8%; Score 58; DB 2; Length 89;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
|||||  
Db 25 HWSYGLRPG 33

## RESULT 6

RHMSG

gonadoliberin precursor - mouse  
N:Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing  
N:Contains: gonadoliberin; gonadoliberin-associated protein (GAP)  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1993 #sequence\_revision 18-Mar-1997 #text\_change 18-Jun-1999  
C:Accession: A47578  
R:Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolic,  
Science 234, 1366-1371, 1986  
A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible fo  
A:Reference number: A47578; MUID:87069928; PMID:3024317  
A:Accession: A47578  
A:Molecule type: DNA  
A:Residues: 1-90 <MAS>  
A:Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175  
C:Genetics:  
A:Introns: 45/3; 77/3  
C:Function:  
A:Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropin  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:22-31/Product: gonadoliberin #status predicted <GLB>  
F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>  
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 35.8%; Score 58; DB 1; Length 90;  
Best Local Similarity 100.0%; Pred. No. 0.31;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
|||||  
Db 23 HWSYGLRPG 31

## RESULT 7

RHUG

gonadoliberin precursor [validated] - human  
N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing  
N:Contains: gonadoliberin-associated protein (GAP); progonadoliberin  
C:Species: Homo sapiens (man)  
C:Date: 17-Mar-1987 #sequence\_revision 21-Jul-1995 #text\_change 08-Dec-2000  
C:Accession: S05308; A26173; A93342; A90108; A01410; S45718  
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.  
Nucleic Acids Res. 17, 6403-6404, 1989  
A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone ge  
A:Reference number: S05308; MUID:89366682; PMID:2671939  
A:Accession: S05308  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-92 <RAY>

A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956  
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot  
A:Reference number: A94090; MUID:86094338; PMID:2867548  
A:Accession: A26173  
A:Molecule type: mRNA

A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749  
A:Experimental source: hypothalamus  
R:Seeburg, P.H.; Adelman, J.P.  
Nature 311, 666-668, 1984  
A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasin  
A:Reference number: A93342; MUID:85012739; PMID:6090951  
A:Accession: A93342  
A:Molecule type: mRNA  
A:Residues: 1-15, 'S', 17-92 <SEE>  
A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357  
A:Experimental source: placenta  
R:Tan, L.; Rousseau, P.  
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982  
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesis  
A:Reference number: A90108; MUID:83126573; PMID:6760865  
A:Accession: A90108  
A:Molecule type: protein  
A:Residues: 24-33 <TAN>  
A:Experimental source: placental trophoblasts  
R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Am  
FEBS Lett. 346, 203-206, 1994  
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by  
A:Reference number: S45718; MUID:94283597; PMID:8013634  
A:Contents: annotation; degradation pathway of synthetic hormone  
C:Genetics:  
A:Gene: GDB:GNRH; LHRH; GRH  
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760  
A:Map position: 8p21-8p11.2  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: gonadoliberin stimulates pituitary secretion of lutropin and follitrop  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting acti  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-92/Product: progonadoliberin #status predicted <PGN>  
F:24-33/Product: gonadoliberin #status experimental <MAT>  
F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status exper  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from followin

Query Match 35.8%; Score 58; DB 1; Length 92;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
|||||  
Db 25 HWSYGLRPG 33

## RESULT 8

RHRTG

gonadoliberin precursor - rat  
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hor  
N:Contains: gonadoliberin; prolactin release-inhibiting factor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 18-Jun-1999  
C:Accession: A40147; B26173; A48410  
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.  
Mol. Endocrinol. 3, 1257-1262, 1989  
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalam  
A:Reference number: A40147; MUID:89384661; PMID:2476669  
A:Accession: A40147  
A:Molecule type: DNA  
A:Residues: 1-92 <BON>

A:Cross-references: GB:M31670; NID:g204447; PIDN:AAA41264.1; PID:g204448  
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonad  
A:Reference number: A94090; MUID:86094338; PMID:2867548  
A:Accession: B26173  
A:Molecule type: mRNA

```

Query Match      35.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      2 HWSYGLRPG 10
        | | | | | | | |
Db      2 HWSYGLRPG 10

```

Endocrinology 134, 1835-1845, 1994  
A; Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved  
A; Reference number: I51423; MUID: 94185563; PMID: 8137750  
A; Accession: I51423  
A; Status: preliminary; translated from GB/EMBL/DBDJ  
A; Molecule type: DNA  
A; Residues: 1-89 <HAY>

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:30:06 ; Search time 10.9058 Seconds  
(without alignments)  
273.361 Million cell updates/sec

Title: US-09-848-834A-15

Perfect score: 162

Sequence: 1 XHWSYGLRPGSSGSLQYIKANSKFEIGTEL 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	46.9	1315	1 BTCLTN	tentoxilysin (EC 3
2	58	35.8	10	1 RHGG	gonadoliberin - pi
3	58	35.8	10	1 RHSHG	gonadoliberin - sh
4	58	35.8	67	2 I78541	gonadoliberin prec
5	58	35.8	89	2 I51423	gonadoliberin prec
6	58	35.8	90	1 RHMSG	gonadoliberin prec
7	58	35.8	92	1 RHUG	gonadoliberin prec
8	58	35.8	92	1 RHRTG	gonadoliberin prec
9	56	34.6	719	2 T52510	gonadoliberin prec
10	55.5	34.3	66	2 S31029	hypothetical prote
11	55	34.0	123	2 G48677	gene 84 protein -
12	54	33.3	10	1 RHAQ1	Ig heavy chain V-D
13	54	33.3	92	2 I50644	gonadoliberin I -
14	53.5	33.0	374	2 E93361	gonadoliberin I pr
15	52	32.1	98	2 I50739	probable mucronate
16	52	32.1	102	2 PH1491	gonadotropin-relea
17	52	32.1	119	2 PH1518	Ig heavy chain V r
18	52	32.1	119	2 PH1519	Ig heavy chain V r
19	52	32.1	135	2 PH1494	Ig heavy chain V r
20	52	32.1	575	2 T06353	Ig heavy chain V r
21	51	31.5	115	2 C27887	isocitrate lyase (
22	51	31.5	112	2 S38715	Ig kappa chain V r
23	51	31.5	123	2 F48677	Ig kappa chain V r
24	51	31.5	208	2 AG2249	Ig heavy chain V-D
25	50.5	31.2	1494	2 T14355	hypothetical prote
26	50	30.9	80	1 RH4D1S	protein-tyrosine-p
27	50	30.9	120	2 A49043	gonadoliberin I pr
28	50	30.9	224	2 A53143	Ig kappa chain V r
29	50	30.9	249	2 A41497	testis-determining 36K antigen pra -

#### ALIGNMENTS

##### RESULT 1

BTCLTN

tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani

N:Alternate names: tetanus neurotoxin

C:Species: Clostridium tetani

C>Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 03-Jun-2002

C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364

R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel

EMBO J. 5, 2495-2502, 1986

A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with

A:Reference number: A25689; MUID:87053814; PMID:3536478

A:Accession: A25689

A:Molecule type: DNA

A:Residues: 1-1315 <EIS>

A:Cross-references: GB:X04436; NID:940769; PIDN:CAA28033.1; PID:g40770

R:Fairweather, N.F.; Lyness, V.A.

Nucleic Acids Res. 14, 7809-7812, 1986

A:Title: The complete nucleotide sequence of tetanus toxin.

A:Reference number: A25757; MUID:87040747; PMID:3774547

A:Accession: A25757

A:Molecule type: DNA

A:Residues: 1-1315 <FAI>

A:Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:g40774

A:Experimental source: strain CN3911

R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.

J. Bacteriol. 165, 21-27, 1986

A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C.

A:Reference number: A25194; MUID:86085672; PMID:3510187

A:Accession: A25194

A:Molecule type: DNA

A:Residues: 743-1315 <FA2>

A:Cross-references: GB:M12739; NID:9144920; PIDN:AAA23282.1; PID:gl44921

A:Accession: B25194

A:Molecule type: protein

A:Residues: 865-894 <FA3>

R:Watsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.

Infect. Immun. 57, 3588-3593, 1989

A:Title: Isolation, purification, and characterization of fragment B, the NH-2-term

A:Reference number: A60759; MUID:90035436; PMID:2478476

A:Accession: A60759

A:Molecule type: protein

A:Residues: 461-475 <MAT>

R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.

J. Immunol. 142, 394-402, 1989

A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.

A:Reference number: JS0098; MUID:89093918; PMID:2463305

A:Contents: annotation; epitope region

R:Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta

Nature 359, 832-835, 1992

A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by prote

A:Reference number: S27125; MUID:93063293; PMID:1331807

A:Contents: annotation

US-09-862-849-2

Query Match 45.7%; Score 74; DB 9; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00052;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31  
Db 1 QYIKANSKFIGITEL 15

RESULT 15

US-09-785-215-4  
; Sequence 4, Application US/09785215  
; Publication No. US20020187157A1  
; GENERAL INFORMATION:  
; APPLICANT: JENSEN, Martin Roland et al.  
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID  
; FILE REFERENCE: 3631-0107P  
; CURRENT APPLICATION NUMBER: US/09/785,215  
; CURRENT FILING DATE: 2001-02-20  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-09-785-215-4

Query Match 45.7%; Score 74; DB 10; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00052;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31  
Db 1 QYIKANSKFIGITEL 15

Search completed: September 24, 2003, 17:33:09  
Job time : 15.5429 secs

QY 16 LOYIKANSKFIGITEL 31  
:|||||  
Db 822 MOYIKANSKFIGITEL 837

RESULT 11  
US-10-241-596-143  
; Sequence 143, Application US/10241596  
; Publication No. US20030166238A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbiological Research Authority  
; APPLICANT: The Speywood Laboratory Limited  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130003  
; CURRENT APPLICATION NUMBER: US/10/241,596  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: US 09/242,689  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: US 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB 9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB 9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 175  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 143  
; LENGTH: 879  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-10-241-596-143

Query Match 46.9%; Score 76; DB 12; Length 879;  
Best Local Similarity 93.8%; Pred. No. 0.023;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31  
:|||||  
Db 829 MOYIKANSKFIGITEL 844

RESULT 12  
US-10-241-596-147  
; Sequence 147, Application US/10241596  
; Publication No. US20030166238A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbiological Research Authority  
; APPLICANT: The Speywood Laboratory Limited  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130003  
; CURRENT APPLICATION NUMBER: US/10/241,596  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: US 09/242,689  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: US 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB 9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB 9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 175  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 147  
; LENGTH: 887  
; TYPE: PRT

; ORGANISM: Clostridium tetani  
US-10-241-596-147

Query Match 46.9%; Score 76; DB 12; Length 887;  
Best Local Similarity 93.8%; Pred. No. 0.023;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31  
:|||||  
Db 829 MOYIKANSKFIGITEL 844

RESULT 13  
US-10-241-596-141  
; Sequence 141, Application US/10241596  
; Publication No. US20030166238A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbiological Research Authority  
; APPLICANT: The Speywood Laboratory Limited  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130003  
; CURRENT APPLICATION NUMBER: US/10/241,596  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: US 09/242,689  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: US 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB 9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB 9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 175  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 141  
; LENGTH: 1315  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-10-241-596-141

Query Match 46.9%; Score 76; DB 12; Length 1315;  
Best Local Similarity 93.8%; Pred. No. 0.036;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31  
:|||||  
Db 829 MOYIKANSKFIGITEL 844

RESULT 14  
US-09-862-849-2  
; Sequence 2, Application US/09862849  
; Patent No. US20020013274A1  
; GENERAL INFORMATION:  
; APPLICANT: Sudhir Paul  
; APPLICANT: Larry J. Smith  
; APPLICANT: Gennady Gololobov  
; TITLE OF INVENTION: Methods for Identifying Inducers and Inhibitors of Proteolysis  
; FILE REFERENCE: UNMC 63123 DIV  
; CURRENT APPLICATION NUMBER: US/09/862,849  
; CURRENT FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: US 09/046,373  
; PRIOR FILING DATE: 1998-03-23  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Clostridium tetani

; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(37)  
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tentoxylisin)  
US-09-848-834A-14

Query Match 53.7%; Score 87; DB 9; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HWSYGLRPGSSGPSL 16  
|||||  
Db 2 HWSYGLRPGSSGPSL 16

RESULT 8  
US-09-848-834A-18  
; Sequence 18, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 18  
; LENGTH: 50  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human  
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Tet  
; OTHER INFORMATION: anus toxoid precursor (Tentoxylisin) protein linked by a spacer t  
; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: MOD\_RES  
; LOCATION: (50)..(50)  
; OTHER INFORMATION: Amidated glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(37)  
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent  
; OTHER INFORMATION: oxylysin  
; NAME/KEY: PEPTIDE  
; LOCATION: (38)..(41)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (42)..(50)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-18

Query Match 53.7%; Score 87; DB 9; Length 50;  
Best Local Similarity 100.0%; Pred. No. 2.6e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HWSYGLRPGSSGPSL 16

Db 2 HWSYGLRPGSSGPSL 16  
|||||

RESULT 9  
US-09-848-834A-2  
; Sequence 2, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Tetanus bacillus  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(16)  
; OTHER INFORMATION: Amino acid sequence 829-844 of the Tetanus  
; OTHER INFORMATION: Toxoid Precursor (Tentoxylisin)  
US-09-848-834A-2

Query Match 46.9%; Score 76; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 0.00029;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 16 LOYIKANSKFIGITEL 31  
:|||||  
Db 1 MQYIKANSKFIGITEL 16

RESULT 10  
US-10-241-596-145  
; Sequence 145, Application US/10241596  
; Publication No. US20030166238A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbiological Research Authority  
; APPLICANT: The Speywood Laboratory Limited  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130003  
; CURRENT APPLICATION NUMBER: US/10/241,596  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: US 09/242,689  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: US 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB 9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB 9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 175  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 145  
; LENGTH: 872  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-10-241-596-145

Query Match 46.9%; Score 76; DB 12; Length 872;  
Best Local Similarity 93.8%; Pred. No. 0.023;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

US-09-848-834A-16

Query Match 54.9%; Score 89; DB 9; Length 36;  
Best Local Similarity 63.3%; Pred. No. 9.5e-06;  
Matches 19; Conservative 2; Mismatches 3; Indels 6; Gaps 1;

QY 2 HWSYGLRPGSSGSPSL-----QYIKANSKF 25  
|||||  
Db 2 HWSYGLRPGSSGSPSLDEKKIAKMERASSVF 31

RESULT 5

US-09-848-834A-17

; Sequence 17, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:

; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17

; LENGTH: 47  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn  
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of  
; OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid seq  
; OTHER INFORMATION: uence 2-10 of the GnRH hormone  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: MOD\_RES  
; LOCATION: (47)..(47)

; OTHER INFORMATION: Amidated-glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(18)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (19)..(34)  
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,  
; NAME/KEY: PEPTIDE  
; LOCATION: (35)..(38)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (39)..(47)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone

US-09-848-834A-17

Query Match 54.9%; Score 89; DB 9; Length 47;  
Best Local Similarity 83.3%; Pred. No. 1.3e-05;  
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSGSPSLQYI 19  
|||||  
Db 2 HWSYGLRPGSSGSPSLKLL 19

RESULT 6

US-09-848-834A-20

; Sequence 20, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:

; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens

; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 20  
; LENGTH: 51  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of h  
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Pl  
; OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: MOD\_RES  
; LOCATION: (51)..(51)

; OTHER INFORMATION: Amidated glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(36)  
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum  
; OTHER INFORMATION: circumsporozoite (CSP) protein  
; NAME/KEY: PEPTIDE  
; LOCATION: (37)..(42)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (43)..(51)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone

US-09-848-834A-20

Query Match 54.9%; Score 89; DB 9; Length 51;  
Best Local Similarity 63.3%; Pred. No. 1.4e-05;  
Matches 19; Conservative 2; Mismatches 3; Indels 6; Gaps 1;

QY 2 HWSYGLRPGSSGSPSL-----QYIKANSKF 25  
|||||  
Db 2 HWSYGLRPGSSGSPSLDEKKIAKMERASSVF 31

RESULT 7

US-09-848-834A-14

; Sequence 14, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:

; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 14  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of tLe  
; OTHER INFORMATION: nRH hormone linked by a spacer to amino acid sequence 947-96  
; OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylisin)  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

Best Local Similarity 100.0%; Pred. No. 3.8e-16;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 HWSYGLRPGSGPSLQYIKANSKFIGITEL 31  
Db 2 HWSYGLRPGSGPSLQYIKANSKFIGITEL 31

## RESULT 2

US-09-848-834A-19  
; Sequence 19, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19  
; LENGTH: 46  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human  
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetanus  
; OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino acid  
; OTHER INFORMATION: sequence 1-10 of GnRH  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: MOD\_RES  
; LOCATION: (46)..(46)  
; OTHER INFORMATION: Amidated glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(31)  
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor  
; OTHER INFORMATION: (Tentoxylisin)  
; NAME/KEY: PEPTIDE  
; LOCATION: (32)..(37)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (38)..(46)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-19

Query Match 99.4%; Score 161; DB 9; Length 46;  
Best Local Similarity 100.0%; Pred. No. 5.9e-16;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSGPSLQYIKANSKFIGITEL 31  
Db 2 HWSYGLRPGSGPSLQYIKANSKFIGITEL 31

## RESULT 3

US-09-848-834A-13  
; Sequence 13, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A

; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13  
; LENGTH: 34  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the  
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 o  
; OTHER INFORMATION: he Measles virus fusion protein,  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(18)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (19)..(34)  
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles  
; OTHER INFORMATION: virus fusion protein, F  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
US-09-848-834A-13

Query Match 54.9%; Score 89; DB 9; Length 34;  
Best Local Similarity 83.3%; Pred. No. 9e-06;  
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSGPSLQYI 19  
Db 2 HWSYGLRPGSGPSLKLL 19

## RESULT 4

US-09-848-834A-16  
; Sequence 16, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the  
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 378-398 o  
; OTHER INFORMATION: the Plasmodium falciparum circumsporozoite (CSP) protein  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(36)  
; OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria  
; OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite  
; OTHER INFORMATION: (CSP) protein

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:57:10 ; Search time 15.5429 Seconds  
(without alignments)  
301.778 Million cell updates/sec

Title: US-09-848-834A-15

Perfect score: 162

Sequence: 1 XHWSYGLRPGSGPSLQYKANSKFIGITEL 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	161	99.4	31	US-09-848-834A-15	Sequence 15, Appl
2	161	99.4	46	US-09-848-834A-19	Sequence 19, Appl
3	89	54.9	34	US-09-848-834A-13	Sequence 13, Appl
4	89	54.9	36	US-09-848-834A-16	Sequence 16, Appl
5	89	54.9	47	US-09-848-834A-17	Sequence 17, Appl
6	89	54.9	51	US-09-848-834A-20	Sequence 20, Appl
7	87	53.7	37	US-09-848-834A-14	Sequence 14, Appl
8	87	53.7	50	US-09-848-834A-18	Sequence 18, Appl
9	76	46.9	16	US-09-848-834A-2	Sequence 2, Appl
10	76	46.9	872	US-10-241-596-145	Sequence 145, App
11	76	46.9	879	US-10-241-596-143	Sequence 143, App
12	76	46.9	887	US-10-241-596-147	Sequence 147, App
13	76	46.9	1315	US-10-241-596-141	Sequence 141, App
14	74	45.7	15	US-09-862-849-2	Sequence 2, Appl
15	74	45.7	15	US-09-785-215-4	Sequence 4, Appl

16	74	45.7	15	12	US-10-237-656-13	Sequence 13, Appl
17	74	45.7	15	12	US-09-405-986-1	Sequence 1, Appl
18	74	45.7	15	12	US-10-223-809A-4	Sequence 4, Appl
19	74	45.7	15	12	US-10-261-446-19	Sequence 19, Appl
20	74	45.7	15	12	US-10-239-313A-618	Sequence 618, App
21	74	45.7	15	15	US-10-204-362-4	Sequence 4, Appl
22	74	45.7	15	15	US-10-223-711-7	Sequence 7, Appl
23	74	45.7	16	12	US-09-930-915A-64	Sequence 64, Appl
24	74	45.7	17	11	US-09-865-294-3	Sequence 3, Appl
25	74	45.7	17	12	US-10-239-313A-619	Sequence 619, App
26	74	45.7	19	12	US-10-239-313A-620	Sequence 620, App
27	74	45.7	27	12	US-10-076-674-7	Sequence 7, Appl
28	74	45.7	28	9	US-09-848-834A-11	Sequence 11, Appl
29	74	45.7	31	12	US-10-237-656-1	Sequence 1, Appl
30	74	45.7	32	12	US-10-237-656-5	Sequence 5, Appl
31	74	45.7	50	9	US-09-943-548-8	Sequence 8, Appl
32	74	45.7	50	15	US-10-339-522-8	Sequence 8, Appl
33	74	45.7	137	12	US-10-237-656-3	Sequence 3, Appl
34	72	44.4	29	15	US-10-238-607-33	Sequence 33, Appl
35	72	44.4	29	15	US-10-238-607-34	Sequence 34, Appl
36	72	44.4	31	10	US-09-983-019-3	Sequence 3, Appl
37	72	44.4	31	10	US-09-983-019-5	Sequence 5, Appl
38	72	44.4	50	9	US-09-943-548-9	Sequence 9, Appl
39	72	44.4	50	15	US-10-339-522-9	Sequence 9, Appl
40	71	43.8	17	11	US-09-305-924-4	Sequence 4, Appl
41	70	43.2	14	12	US-09-932-165-1481	Sequence 1481, Ap
42	70	43.2	14	12	US-10-116-118-32	Sequence 32, Appl
43	70	43.2	14	12	US-09-935-384-709	Sequence 709, App
44	70	43.2	14	12	US-10-062-109A-760	Sequence 760, App
45	70	43.2	14	12	US-09-942-052-710	Sequence 710, App

#### ALIGNMENTS

#### RESULT 1

US-09-848-834A-15  
; Sequence 15, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15  
; LENGTH: 31  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the  
; OTHER INFORMATION: RH hormone linked by a spacer to amino sequence 830-844 of the  
; OTHER INFORMATION: tanus toxoid precursor (Tentoxylisin)  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(31)  
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor  
; OTHER INFORMATION: (tentoxylisin)  
US-09-848-834A-15

Query Match 99.4%; Score 161; DB 9; Length 31;



PN WO9310152-A1.  
 XX  
 PD 27-MAY-1993.  
 XX  
 PF 11-NOV-1992; 92WO-EP02591.  
 XX  
 PR 16-NOV-1991; 91GB-0024390.  
 PR 27-FEB-1992; 92US-0842694.  
 XX  
 XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PA  
 XX Cohen J, De Wilde M;  
 PI WPI; 1993-182494/22.  
 DR N-PSDB; AAQ42567.  
 XX  
 XX Hybrid protein comprising Plasmodium circumsporozoite protein and  
 PT HbsAg - useful as a vaccine for treating patients susceptible to  
 PT Plasmodium infections  
 XX  
 PS Disclosure; Fig 9; 59pp; English.  
 XX  
 CC This sequence represents the RTS\* hybrid protein which is encoded by  
 CC the RTS\* expression cassette. This hybrid consists of a methionine  
 CC residue derived from S. cerevisiae TDH3 gene sequence, three amino  
 CC acids, Met-Ala-Pro, derived from a nucleotide sequence created by  
 CC the cloning procedure used to construct the hybrid gene, a stretch  
 CC of 189 amino acids representing amino acids 210 to 398 of the  
 CC circumsporozoite protein (CSP) of Plasmodium falciparum strain NF54,  
 CC an amino acid Arg created by the cloning procedure, four amino acids,  
 CC Pro-Val-Thr-Asn, representing the four carboxy terminal residues of  
 CC Hepatitis B virus (HBV), adw serotype, preS2 protein, and a stretch  
 CC of 226 amino acids specifying the S protein of HBV, adw serotype.  
 CC This protein, and RTS (see also AAR37796), may be combined with an  
 CC adjuvant and used in a vaccine for preventing plasmodium infections.  
 CC The vaccines produce a humoral response and also a cellular immune  
 CC response.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 424 AA;  
 SQ  
 Query Match 49.7%; Score 91; DB 14; Length 424;  
 Best Local Similarity 55.3%; Pred. No. 0.00011;  
 Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;  
 QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNS 36  
 Db ::|||:| |||| |||| |||| |||| ||||  
 152 IRPGSANKPKDLDYANDIERKICKMEKSCSVFNVNS 189  
 RESULT 15  
 AAG63663  
 ID AAG63663 standard; peptide; 33 AA.  
 XX  
 AC AAG63663;  
 XX  
 DT 29-OCT-2001 (first entry)  
 XX  
 DE Peptide comprising conjugation sites for a pseudopeptide.  
 XX  
 KW Pseudopeptide; immunomodulation; adjuvant; antigen presenting cell;  
 KW macrophage; dendritic cell; vaccine; autoimmune disease.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200146127-A1.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 22-DEC-1999; 99WO-IB02038.  
 XX  
 PR 22-DEC-1999; 99WO-IB02038.  
 XX  
 XX

PA (OMPH-) OM-PHARMA.  
 XX  
 PI Bauer J, Martin OR, Rodriguez S;  
 XX  
 DR WPI; 2001-502469/55.  
 XX  
 PT New amphiphilic acylated pseudopeptides having a functionalized  
 PT auxiliary spacer, useful as immunomodulators, e.g. as adjuvants in  
 PT vaccines  
 XX  
 XX Example 3; Page 61; 166pp; French.  
 PS  
 XX The specification describes N-Acylated pseudopeptides, which have  
 XX a neutral or charged acidic group at one terminal and a functionalized  
 CC auxiliary spacer at the other. The pseudopeptides show immunomodulatory  
 CC and adjuvant action, based on activation of antigen presenting cells  
 CC (e.g. macrophages or dendritic cells), induction of differentiation of  
 CC dendritic cells, induction of cytokine production and induction of  
 CC maturation of immunocompetent cell strains originating from hematopoietic  
 CC and lymphoid organs. They reinforce humoral and cellular immunity. They  
 CC can be grafted onto antigens (to modulate immune response) or onto drugs  
 CC (to improve the therapeutic activity or targeting). The pseudopeptides  
 CC are thus useful in human or veterinary medicine as immunizing or  
 CC diagnostic agents. Typically, they are used as adjuvants together with  
 CC (or covalently bonded to) antigens for vaccination against viral,  
 CC parasitic/protozoal, microbial or fungal infections; incubated with blood  
 CC cells ex vivo, to render the cells immunocompetent before reintroduction  
 CC in vivo; or used in therapy of certain autoimmune diseases. The  
 CC pseudopeptides are useful as carriers for antigens or other therapeutic  
 CC agents due to their ability to form non-covalent bonds via the  
 CC hydrophobic or hydrophilic auxiliary spacer. The present sequence  
 CC represents a peptide, which has conjugation sites for the pseudopeptides  
 CC of the invention.  
 XX  
 SQ Sequence 33 AA;  
 Query Match 48.9%; Score 89.5; DB 22; Length 33;  
 Best Local Similarity 76.9%; Pred. No. 9e-06;  
 Matches 20; Conservative 3; Mismatches 2; Indels 1; Gaps 1;  
 QY 11 SSGPSLDEKKIAKMEKASSVFNVNS 36  
 Db ::|:||||||| ||||| ||||| ||||| |||||  
 9 NANPDI-EKKIAMEKASSVFNVNS 33  
 Search completed: September 24, 2003, 17:40:07  
 Job time : 41.687 secs

PA (USDC ) US SEC OF COMMERCE.  
 PA (USGO ) US GOVERNMENT.  
 PA (USSA ) US SEC OF ARMY.

XX McCutchan TF, Dame JB, Williams JL, Schneider I;  
 XX

XX WPI; 1986-008635/02.  
 DR N-PSDB; AAN60362.

XX New immunologically active pure synthetic peptide(s) - used for  
 PT protection against infection by malaria parasite.

XX Disclosure; Fig 2; 49pp; English.

XX The Plasmodium CS gene was used to isolate peptides capable of  
 CC inducing an immune response to the parasite. Peptide antigens may  
 CC be synthesised in pure form and used to generate an immune  
 CC response in vaccination against malaria. The featured repeat  
 CC units are claimed and must be present in copies of 2-1000.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX

SQ Sequence 412 AA;

Query Match 49.7%; Score 91; DB 7; Length 412;  
 Best Local Similarity 55.3%; Pred. No. 0.00011;  
 Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPGSSGPSLD-----EKKIAKMEKASSVFNVNS 36  
 :|||: | |||| |||| |||| ||||  
 Db 361 IKPGSANKPKDELNDYENDIEKKICKMEKCSVFNVNS 398

RESULT 13  
 AAR37796  
 ID AAR37796 standard; Protein; 424 AA.

XX AAR37796;

DT 25-MAR-2003 (updated)  
 DT 27-SEP-1993 (first entry)

XX RTS protein.

XX RTS; expression cassette; hybrid protein; S. cerevisiae; TDH3;  
 KW cloning; circumsporozoite protien; CSP; Plasmodium falciparum;  
 KW strain 7G8; hepatitis B virus; HBV; adw serotype; pres2 protein;  
 KW S protein.

XX Synthetic.

XX Key Location/Qualifiers

FT Region 1 /note= "Derived from S. cerevisiae TDH3 gene sequence"

FT Region 2..4 /note= "Cloning artefact"

FT Protein 5..193 /note= "Represents amino acids 210-398 of the CSP of

FT P. falciparum"

FT Region 194..197

FT Protein /note= "Carboxy terminal amino acids from HBV (adw  
 serotype) pres2 protein"

FT /note= "S protein of HBV (adw serotype)"

XX WO9310152-A1.

XX 27-MAY-1993.

XX 11-NOV-1992; 92WO-EF02591.

XX 16-NOV-1991; 91GB-0024390.

XX 27-FEB-1992; 92US-0842694.

XX

PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX Cohen J, De Wilde M;

XX WPI; 1993-182494/22.

DR N-PSDB; AAQ42566.

XX Hybrid protein comprising Plasmodium circumsporozoite protein and  
 PT HfsAg - useful as a vaccine for treating patients susceptible to  
 PT Plasmodium infections

XX Disclosure; Fig 5; 59pp; English.

XX This sequence represents the RTS hybrid protein which is encoded by  
 CC the RTS expression cassette. This hybrid consists of a methionine  
 CC residue derived from S. cerevisiae TDH3 gene sequence, three amino  
 CC acids, Met-Ala-Pro, derived from a nucleotide sequence created by  
 CC the cloning procedure used to construct the hybrid gene, a stretch  
 CC of 189 amino acids representing amino acids 210 to 398 of the  
 CC circumsporozoite protien (CSP) of Plasmodium falciparum strain 7G8,  
 CC an amino acid Arg created by the cloning procedure, four amino acids,  
 CC Pro-Val-Thr-Asn, representing the four carboxy terminal residues of  
 CC hepatitis B virus (HBV), adw serotype, pres2 protein, and a stretch  
 CC of 226 amino acids specifying the S protein of HBV, adw serotype.  
 CC This protein, and RTS\* (see also AAR37797), may be combined with an  
 CC adjuvant and used in a vaccine for preventing plasmodium infections.  
 CC The vaccines produce a humoral response and also a cellular immune  
 CC response.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX

SQ Sequence 424 AA;

Query Match 49.7%; Score 91; DB 14; Length 424;

Best Local Similarity 55.3%; Pred. No. 0.00011;

Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPGSSGPSLD-----EKKIAKMEKASSVFNVNS 36  
 :|||: | |||| |||| |||| ||||  
 Db 156 IKPGSANKPKDELNDYENDIEKKICKMEKCSVFNVNS 193

RESULT 14

AAR37797  
 ID AAR37797 standard; Protein; 424 AA.

XX AAR37797;

DT 25-MAR-2003 (updated)

DT 27-SEP-1993 (first entry)

XX RTS\* protein.

XX RTS; expression cassette; hybrid protein; S. cerevisiae; TDH3;  
 KW cloning; circumsporozoite protien; CSP; Plasmodium falciparum;  
 KW strain 7G8; hepatitis B virus; HBV; adw serotype; pres2 protein;  
 KW S protein.

XX Synthetic.

XX Key Location/Qualifiers

FT Region 1 /note= "Derived from S. cerevisiae TDH3 gene sequence"

FT Region 2..4 /note= "Cloning artefact"

FT Protein 5..193 /note= "Represents amino acids 210-398 of the CSP of

FT P. falciparum"

FT Region 194..197

FT Protein /note= "Carboxy terminal amino acids from HBV (adw  
 serotype) pres2 protein"

FT /note= "S protein of HBV (adw serotype)"

XX



The polypeptide is prepd. by genetic engineering of genes encoding  
the P. falciparum circumsporozoite (CS) protein [Dame et al.,  
Science 225 : 593 (1984)], and the influenza virus non-structural  
protein 1 (NS1). [Baez et al., Nucleic Acids Research, 8 : 5845  
(1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1  
(NS1-81) is linked to a synthetic sequence encoding four repeat  
units from the immunodominant region, which in turn is linked via  
a synthetic sequence to DNA encoding Region I contg. flanking  
region less the 18 AA signal region. This is linked to DNA  
encoding Region II-contg. flanking region. The Pro residue sep-  
arating the Asp (at the C-terminal of the linker) from the Region  
I-contg. CS flanking region is an artifact of a filled-in BamHI  
site; the Gly separating the Region I and II-contg. CS flanking  
regions is an artifact of a synthetic FokI/TthIII I linker. The  
peptide can be used in a vaccine for protection against malaria.  
See also AAR12306-R12311 and AAR13175-R13179.  
(Updated on 25-MAR-2003 to correct PA field.)

xx Sequence 335 AA;

SQ Query Match 49.7%; Score 91; DB 12; Length 335;  
Best Local Similarity 55.3%; Pred. No. 8.5e-05;  
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKTKAKMEKASSVENVVNS 36  
::|||: | ||| |||| |||||||||

Db 284 IKPGSANKPKDELVDYENDIEKKICKMEKCSSVENVVNS 321

RESULT 10

AAR13179

ID ID AAR13179 standard; Protein; 335 AA.

XX AC AAR13179;

XX DT 25-MAR-2003 (updated)

XX DT 29-AUG-1991 (first entry)

XX XX NS1\_81(NVDP)4RufAuth.

XX DE

XX KW Immunogenic determinant; circumporozoite; CS; vaccine; malaria;

XX KW hybrid; influenza virus; non-structural protein 1; fusion.

XX OS Plasmodium falciparum.

OS Influenza virus (A/PR/8/34).

FH Key Location/Qualifiers

FT FT Region 1..81 /label= N-terminal of NS1

FT FT /note= "Influenza virus nonstructural protein 1"

FT FT Region 82..97

FT FT /label= immunodominant repeat region

FT FT /note= "four variant tetrapeptide repeat units"

FT FT Peptide 98..103

FT FT /label= synthetic linker

FT FT Region 104

FT FT /label= artifact

FT FT /note= "see comments"

FT FT Region 105..209

FT FT /label= AAS 19-123 of CS protein

FT FT /note= "Region I contg. flanking region less  
signal sequence"

FT FT Region 210

FT FT /label= artifact

FT FT /note= "see comments"

FT FT Region 211..335

FT FT /label= AAS 288-412 of CS protein

FT FT /note= "Region II flanking region"

XX EP432965-A.

PN

XX PD 19-JUN-1991.

YY

```

FT Region 88
FT /label= artifact
FT /note= "see comments"
FT 89..193
FT /label= AAs 19-123 of CS protein
FT /note= "Region I contg. flanking region less
FT signal sequence"
FT 194
FT /label= artifact
FT /note= "see comments"
FT 195..319
FT /label= AAs 288-412 of CS protein
FT /note= "Region II flanking region"
PN EP432965-A..
XX 19-JUN-1991.
PD 06-DEC-1990; 90EP-0313257.
XX 08-DEC-1989; 89US-0447746.
XX (SMIK ) SMITHKLINE BEECHAM.
PA (USSA ) US SEC OF ARMY.
PA (BIOM-) BIOMEDICAL RES INST.
PA (GROSS/) GROSS M S.
XX
PI Gross MS, Gordon DM, Hollingdale MR;
XX WPI; 1991-179771/25.
XX
DR Polypeptide comprising immunogenic determinants from P falciparum
XX - for vaccine against malaria infection in humans.
XX
PS Example 2; Page 10; 18pp; English.
XX
CC The polypeptide is prepd. by genetic engineering of genes encoding
CC the P. falciparum circumsporozoite (CS) protein [Dame et al.,
CC Science 225 : 593 (1984)], and the influenza virus non-structural
CC protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
CC (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
CC I contg. flanking regionless the 18 AA signal region, which in
CC turn is fused to DNA encoding Region II-contg. flanking region.
CC This CS fusion is designated RlfAuth. The pro residue separating
CC the Asp (at the C-terminal of the linker) from RlfAuth is an arti-
CC fact of a filled in BamHI site; the Gly separating Region I and
CC Region II-contg. CS flanking regions is an artifact of a synthetic
CC FokI/NotIII I linker. The peptide can be used in a vaccine for
CC protection against malaria.
CC The complete nucleotide and AA sequences are given in EP-304720,
CC filed May 1, 1990.
CC See also AAR12306-R12311 and AAR13175-R13179.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 319 AA;
XX
Query Match 49.7%; Score 91; DB 12; Length 319;
Best Local Similarity 55.3%; Pred. NO. 8e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
QY 7 LRPSSGSPSLD-----EKIKAKMEKASSVFNVNVS 36
DB 268 IKPSANKPKDELVDYNDIEKKICKMEKCSVFNVNVS 305
:||||: | |||| |||| |||| |||| ||||
RESULT 8
AAR13177
ID AAR13177 standard; Protein; 327 AA.
XX
AC AAR13177;
XX
DT 25-MAR-2003 (updated)
XX

```

```

DT 29-AUG-1991 (first entry)
XX
XX NS1_81-RLfAuth + (NANP)2.
XX
KW Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
KW hybrid; influenza virus; non-structural protein 1; fusion.
XX
XX Plasmodium falciparum.
XX Influenza virus (A/PR/8/34/).
XX
FH Key Location/Qualifiers
FT Region 1..81
FT /label= N-terminal of NS1
FT /note= "Influenza virus nonstructural protein 1"
FT Peptide 82..87
FT /label= synthetic linker
FT Region 88
FT /label= artifact
FT /note= "see comments"
FT Region 89..193
FT /label= AAs 19-123 of CS protein
FT /note= "Region I contg. flanking region less
FT signal sequence"
FT Region 194..201
FT /label= immunodominant repeat region
FT /note= "two tetrapeptide repeat units"
FT Region 202
FT /label= artifact
FT /note= "see comments"
FT Region 203..327
FT /label= AAs 288-412 of CS protein
FT /note= "Region II flanking region"
XX
PN EP432965-A.
XX
XX 19-JUN-1991.
XX
XX 06-DEC-1990; 90EP-0313257.
XX
XX 08-DEC-1989; 89US-0447746.
XX (SMIK ) SMITHKLINE BEECHAM.
PA (USSA ) US SEC OF ARMY.
PA (BIOM-) BIOMEDICAL RES INST.
PA (GROSS/) GROSS M S.
XX
XX Gross MS, Gordon DM, Hollingdale MR;
XX WPI; 1991-179771/25.
XX
XX Polypeptide comprising immunogenic determinants from P falciparum
XX - for vaccine against malaria infection in humans.
XX
PS Example 3; Page 10; 18pp; English.
XX
CC The polypeptide is prepd. by genetic engineering of genes encoding
CC the P. falciparum circumsporozoite (CS) protein [Dame et al.,
CC Science 225 : 593 (1984)], and the influenza virus non-structural
CC protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
CC (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
CC (NS1_81) is linked via a synthetic sequence to DNA encoding Region
CC I contg. flanking region less the 18 AA signal region. This is
CC linked to a synthetic sequence encoding two repeat units from the
CC immunodominant region, which in turn is fused to DNA encoding the
CC Region II-contg. flanking region. The pro residue separating the
CC Asp (at the C-terminal of the linker) from the Region I-contg. CS
CC flanking region is an artifact of a filled-in BamHI site; the Gly
CC separating the repeat units and the Region II-contg. CS flanking
CC region is an artifact of a synthetic FokI/NotIII I linker. The
CC peptide can be used in a vaccine for protection against malaria.
CC See also AAR12306-R12311 and AAR13175-R13179.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX

```

```

FT FT /label= artifact
FT FT /note= "see comments"
FT FT 89..193
FT FT /label= AAs 19-123 of CS protein
FT FT /note= "Region I contg. flanking region less
FT FT signal sequence"
FT FT
FT FT 194..309
FT FT /label= AAs 297-412 of CS protein
FT FT /note= "Region II flanking region minus 9 N-term-
FT FT inal AAs"
XX XX
XX XX EP432965-A.
XX XX
XX XX 19-JUN-1991.
XX XX
XX XX 06-DEC-1990; 90EP-0313257.
XX XX
XX XX 08-DEC-1989; 89US-0447746.
XX XX (SMIK ) SMITHKLINE BEECHAM.
XX XX (USSA ) US SEC OF ARMY.
XX XX (BIOM-) BIOMEDICAL RES INST.
XX XX (GROS/) GROSS M S.
XX XX
XX XX Gross MS, Gordon DM, Hollingdale MR;
XX XX WPI; 1991-179771/25.
XX XX
XX XX Polypeptide comprising immunogenic determinants from P falciparum
XX XX - for vaccine against malaria infection in humans.
XX XX
XX XX Example 1; Page 7; 18pp; English.
XX XX
XX XX The polypeptide is prepd. by genetic engineering of genes encoding
XX XX the P. falciparum circumsporozoite (CS) protein [Dame et al.,
XX XX Science 225 : 593 (1984)], and the influenza virus non-structural
XX XX protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
XX XX (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX XX (NS1-81) is linked via a synthetic sequence to DNA encoding Region
XX XX I contg. flanking regionless the 18 AA signal region, which in
XX XX turn is fused to DNA encoding Region II-contg. flanking region
XX XX less the first nine N-terminal AAs. This CS fusion is designated
XX XX Rldelta9. The pro residue separating the Asp (at the C-terminal
XX XX of the linker) from Rldelta9 is an artifact of a filled in BamHI
XX XX site. The peptide can be used in a vaccine for protection against
XX XX malaria.
XX XX See also AAR12306-R12311 and AAR13176-R13179.
XX XX (Updated on 25-MAR-2003 to correct PA field.)
XX XX
XX XX Query Match 49.7%; Score 91; DB 12; Length 309;
XX XX Best Local Similarity 55.3%; Pred. No. 7.7e-05;
XX XX Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
XX XX
XX XX 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVVNS 36
XX XX ::|||: | |||| |||| |||| ||||
XX XX 258 IKPGSANKPRDELVDYENDIEKKICKMEKCSVFNVVNS 295

RESULT 6
AAR07945
ID AAR07945 standard; protein; 319 AA.
XX XX
XX XX AAR07945;
XX XX
XX XX 25-MAR-2003 (updated)
XX XX 22-FEB-1991 (first entry)
XX XX
XX XX NS181RLFAuth plasmid product.
XX XX Malaria; vaccine.
XX XX

```

```

OS OS Plasmodium falciparum.
XX XX
XX XX Location/Qualifiers
XX XX 1..81
XX XX /label= NS181 protein fragment
XX XX /note= "from plasmid pMG-1"
XX XX
XX XX Domain
XX XX 89..193
XX XX /label= Fragment of circumsporozite protein
XX XX 204..319
XX XX /label= Fragment of circumsporozite protein
XX XX
XX XX EP398540-A.
XX XX
XX XX 22-NOV-1990.
XX XX
XX XX 01-MAY-1990; 90EP-0304720.
XX XX
XX XX 03-MAY-1989; 89US-0346863.
XX XX (SMIK ) SMITHKLINE BEECHAM.
XX XX (GROS/) GROSS M S.
XX XX
XX XX Gross MS, Young JF;
XX XX
XX XX WPI; 1990-350299/47.
XX XX N-ESDB; AAQ06580.
XX XX
XX XX New polypeptide used in malaria vaccine - comprises immunogenic
XX XX determinants from 1st and 2nd flanking regions of plasmodium
XX XX surface protein and intermediate repeat domain
XX XX
XX XX Example 2; Page 11-12; 24pp; English.
XX XX
XX XX The product is useful in preparation of vaccines for treatment and
XX XX prophylaxis of plasmodium sporozite infection. It may be easily
XX XX produced in large pure quantities from a transformed E.coli
XX XX expression system.
XX XX (Updated on 25-MAR-2003 to correct PA field.)
XX XX
XX XX Query Match 49.7%; Score 91; DB 11; Length 319;
XX XX Best Local Similarity 55.3%; Pred. No. 8e-05;
XX XX Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
XX XX
XX XX 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVVNS 36
XX XX ::|||: | |||| |||| |||| ||||
XX XX 268 IKPGSANKPRDELVDYENDIEKKICKMEKCSVFNVVNS 305

RESULT 7
AAR13176
ID AAR13176 standard; Protein; 319 AA.
XX XX
XX XX AAR13176;
XX XX
XX XX 25-MAR-2003 (updated)
XX XX 29-AUG-1991 (first entry)
XX XX
XX XX NS1_81-RLFAuth.
XX XX
XX XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
XX XX hybrid; Influenza virus; non-structural protein 1; fusion.
XX XX
XX XX Plasmodium falciparum.
XX XX Influenza virus (A/PR/8/34).
XX XX
XX XX Location/Qualifiers
XX XX 1..81
XX XX /label= N-terminal of NS1
XX XX /note= "Influenza virus nonstructural protein 1"
XX XX 82..87
XX XX /label= synthetic linker
XX XX

```

XX WO200185763-A2.  
 PN 15-NOV-2001.  
 PD 04-MAY-2001; 2001WO-US14363.  
 PF 05-MAY-2000; 2000US-202328P.  
 XX (APHT-) APHTON CORP.  
 XX Grimes S, Michaeli D, Stevens VC;  
 PI WPI; 2002-049440/06.  
 DR Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue -  
 XX Disclosure; Page 28; 43pp; English.  
 XX The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hyper trophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a peptide from circumsporozoite protein, CSP, a promiscuous helper  
 CC T-cell peptide epitope used in the immunogen of the invention.  
 XX Sequence 20 AA;  
 SQ

Query Match 51.9%; Score 95; DB 23; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DEKKIAKMEKASSVFNVVNS 36  
 Db 1 DEKKIAKMEKASSVFNVVNS 20  
 |||||

RESULT 4  
 AAU11423  
 ID AAU11423 standard; peptide; 33 AA.  
 XX  
 AC AAU11423;  
 XX  
 DT 12-MAR-2002 (first entry)  
 DE Synthetic immunogen peptide 4.  
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX Chimeric - Plasmodium falciparum.  
 OS Chimeric - Mammalia.  
 OS Synthetic.  
 XX Key Location/Qualifiers  
 FH Key 1..20  
 FT Peptide /note= "Malaria CSP protein (378-398 aa)"  
 FT Peptide 21..24  
 FT Peptide /note= "Spacer peptide"  
 FT Peptide 25..33

FT Modified-site /note= "Gonadotropin releasing hormone epitope"  
 FT 33  
 FT /note= "Amidated glycine or glycineamide"  
 XX WO200185763-A2.  
 PN 15-NOV-2001.  
 XX 04-MAY-2001; 2001WO-US14363.  
 XX 05-MAY-2000; 2000US-202328P.  
 XX (APHT-) APHTON CORP.  
 XX Grimes S, Michaeli D, Stevens VC;  
 PI WPI; 2002-049440/06.  
 DR Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue -  
 XX Claim 11; Page 8; 43pp; English.  
 XX The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hyper trophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX Sequence 33 AA;  
 SQ

Query Match 51.9%; Score 95; DB 23; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DEKKIAKMEKASSVFNVVNS 36  
 Db 1 DEKKIAKMEKASSVFNVVNS 20  
 |||||

RESULT 5  
 AAR13175  
 ID AAR13175 standard; Protein; 309 AA.  
 XX  
 AC AAR13175;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 29-AUG-1991 (first entry)  
 XX NS1\_81-RLfdelta9.  
 DE Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;  
 KW hybrid; influenza virus; non-structural protein 1; fusion.  
 XX Plasmodium falciparum.  
 OS Influenza virus (A/PR/8/34/).  
 XX Key Location/Qualifiers  
 FH Key 1..81  
 FT Region /label= N-terminal of NS1  
 FT Peptide /note= "Influenza virus nonstructural protein 1"  
 FT 82..87  
 FT Peptide /label= synthetic linker  
 FT 88



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:58:36 ; Search time 40.687 Seconds  
(without alignments)  
140.442 Million cell updates/sec

Title: US-09-848-834A-16  
Perfect score: 183  
Sequence: 1 XHWSYGLRPGSSGSLDEKKIAKMEKASSVFNVNS 36

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	182	99.5	36	AAU11427	Synthetic immunogen
2	182	99.5	51	AAU11431	Synthetic immunogen
3	95	51.9	20	AAU11414	P. falciparum circ
4	95	51.9	33	AAU11423	Synthetic immunogen
5	91	49.7	309	AAU13175	NS1_81-Rlfdelta9.
6	91	49.7	319	AAU07945	NS18IRLAuth plasm
7	91	49.7	319	AAU13176	NS1_81-RlfAuth. P
8	91	49.7	327	AAU13177	NS1_81-RlfAuth + (
9	91	49.7	335	AAU13178	NS1_81(NANP)4RLFAU

10	91	49.7	335	12	AAU13179	NS1_81(NVDP)4RLFAU
11	91	49.7	411	9	AAU13144	Sequence encoded b
12	91	49.7	412	7	AAU60416	CS protein of mala
13	91	49.7	424	14	AAU37796	RIS protein. Syn
14	91	49.7	424	14	AAU37797	RIS* protein. Syn
15	89.5	48.9	33	22	AAU63663	Peptide comprising
16	89.5	48.9	33	22	AAU63516	A peptide which ma
17	89	48.6	19	22	AAU98951	Vaccine related MH
18	89	48.6	21	10	AAU91504	Sequence of modifi
19	89	48.6	21	16	AAU82586	Plasmodium falcipa
20	89	48.6	21	16	AAU78920	Malaria circumspor
21	89	48.6	21	16	AAU75955	P. falciparum CS p
22	89	48.6	21	16	AAU70912	Malaria circumspor
23	89	48.6	21	17	AAU05612	Circumsporozoite h
24	89	48.6	21	18	AAU33440	T-cell stimulatory
25	89	48.6	21	20	AAU23252	Peptide derived fr
26	89	48.6	21	21	AAU80071	Pathogen derived T
27	89	48.6	21	21	AAU54553	T helper cell (Th)
28	89	48.6	21	21	AAU58777	Unidentified pepti
29	89	48.6	21	22	AAU99706	Plasmodium falcipa
30	89	48.6	21	22	AAU62428	Plasmodium falcipa
31	89	48.6	21	22	AAU84517	Plasmodium falcipa
32	89	48.6	21	22	AAU88269	Plasmodium falcipa
33	89	48.6	21	22	AAU89366	Plasmodium falcipa
34	89	48.6	21	22	AAU84447	Sequence of T help
35	89	48.6	21	22	AAU98457	Plasmodium falcipa
36	89	48.6	21	23	AAU16181	Zinc transporter p
37	89	48.6	21	23	AAU11372	P falciparum HLA b
38	89	48.6	21	23	AAU05781	P falciparum circ
39	89	48.6	21	23	AAU78050	Loosely MHC-restri
40	89	48.6	21	23	AAU01952	158P107 related HL
41	89	48.6	21	23	AAU51501	Malaria circumspor
42	89	48.6	21	23	AAU34858	P. falciparum circ
43	89	48.6	21	23	AAU91545	P. falciparum circ
44	89	48.6	21	23	AAU95369	Plasmodium circum
45	89	48.6	21	23	AAU94469	Plasmodium falcipa

# ALIGNMENTS

RESULT 1  
AAU11427 standard; peptide; 36 AA.  
XX AAU11427;  
AC AAU11427;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Synthetic immunogen peptide 8.

XX Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
luteinising hormone releasing hormone; LHRH; contraceptive;  
promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Chimeric - Plasmodium falciparum.  
OS Chimeric - Mammalia.  
OS Synthetic.

Key	Location/Qualifiers
Misc-difference 1	/label= OTHER
FT	/note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT	1..10
FT	Peptide
FT	/note= "Gonadotrophin releasing hormone epitope"
FT	11..16
FT	Peptide
FT	/note= "Spacer peptide"
FT	17..36
FT	Peptide
FT	/note= "Malaria CSP protein (378-398 aa)"
XX	WO200185763-A2.

Query Match 31.5%; Score 51; DB 16; Length 208;  
Best Local Similarity 42.9%; Pred. No. 23;  
Matches 9; Conservative 5; Mismatches 7; Indels

7 LRPSSGSPSLQYIKANSKFIG 27  
| : | | | | : | : | : |  
52 LKPGSQGPDPVQTLTQLKQLG 72

Search completed: September 24, 2003, 17:47:20  
 Job time : 31.0554 secs

```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA Nabissi M.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC EMBL; AF046801; AAD02427.1; -.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRNI.
DR PROSITE; PS00473; GnRH; 1.
KW Amidation; Hormone.
FT NON_TER 1
FT NON_TER 87
SQ SEQUENCE 87 AA; 9871 MW; OD2463533D96782A CRC64;

Query Match 32.1%; Score 52; DB 13; Length 87;
Best Local Similarity 88.9%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db ||||| ||
22 HWSYGLSPG 30

RESULT 13
Q8JFY3 PRELIMINARY; PRT; 94 AA.
AC Q8JFY3;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE GnRH prohormone precursor.
GN GNRH1.
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Farahmand H., Rahman M.A., Sohn F., Hwang G.-L., Maclean N.;
RT "Isolation and Expression of Tilapia (Oreochromis niloticus) Serine 8
RT Type GnRH Coding and Regulatory Sequences.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467291; AAM0220.1; -.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRNI.
DR PROSITE; PS00473; GnRH; 1.
KW Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 32 GONADOTROPIN-RELEASING HORMONE.
FT CHAIN 36 94 GNRH-ASSOCIATED PEPTIDE.
SQ SEQUENCE 94 AA; 10396 MW; E57DBA832FC078D7 CRC64;

Query Match 32.1%; Score 52; DB 13; Length 94;
Best Local Similarity 88.9%; Pred. No. 6.9;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db ||||| ||
24 HWSYGLSPG 32

```

```

RESULT 14
Q8UW80 PRELIMINARY; PRT; 96 AA.
ID Q8UW80;
AC Q8UW80;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Seabream-type gonadotropin-releasing hormone precursor (Gonadoliberin)
DE (GnRH) (LH-RH) (Luliberlin).
OS Verasper moseri (Barfin flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Pleuronectidae; Verasper.
OX NCBI_TaxID=98923;
RN [1]
RP SEQUENCE FROM N.A.
RA Amano M.;
RT "Molecular cloning of three cDNAs encoding GnRH in the brain of barfin
RT flounder.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC EMBL; AB066360; BAB83984.1; -.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRNI.
DR PROSITE; PS00473; GnRH; 1.
KW Amidation; Hormone; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 36 SEABREAM-TYPE GONADOTROPIN-RELEASING
HORMONE.
FT CHAIN 40 96 GONADOTROPIN-RELEASING HORMONE ASSOCIATED
PEPTIDE.
SQ SEQUENCE 96 AA; 10560 MW; FA3202565EA00DCC CRC64;

Query Match 32.1%; Score 52; DB 13; Length 96;
Best Local Similarity 88.9%; Pred. No. 7;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db ||||| ||
28 HWSYGLSPG 36

RESULT 15
Q8YR98 PRELIMINARY; PRT; 208 AA.
ID Q8YR98;
AC Q8YR98;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hypothetical protein All3550.
GN ALL3550.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003593; BAB75249.1; -.
DR InterPro; IPR002477; PG_binding.

```

DR EMBL; AP003190; BAB1077.1; -.  
DR InterPro; IPR000175; Na/ntran\_symport.  
DR Pfam; PF00209; SNF; 1.  
DR PRINTS; PR00176; NANEUSMPORT.  
DR ProDom; PD000448; Na/ntran\_symport; 1.  
DR PROSITE; PS0267; NA\_NEUTROTRAN\_SYMP\_3; 1.  
KW Complete proteome.  
SQ SEQUENCE 437 AA; 47945 MW; AB012AC43DA9A813 CRC64;

Query Match 32.4%; Score 52.5; DB 16; Length 437;  
Best Local Similarity 50.08; Pred. No. 32;  
Matches 13; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Oy 3 WSYGLRPGSGPSLQYIKANSKFTGI 28  
Db :||||| ||||| :| | |  
273 FAYGLNP-SSGPSLLFITLPKIFQGM 297

## RESULT 9

Q8JIF3 PRELIMINARY; PRT; 64 AA.  
AC Q8JIF3;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Gonadotropin-releasing hormone (Fragment).  
OS Dentex dentex.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;  
OC Sparidae; Dentex.  
OX NCBI\_TaxID=94951;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kato M., Elmesiry G.E.;  
RT "Sequence comparison of GnRH genes in closely-related Sparidae fishes."  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB089313; BAC07229.1; -.  
DR InterPro; IPR002012; GnRH.  
DR Pfam; PF00446; GnRH; 1.  
DR PRINTS; PR01341; GONADOLIBRN1.  
DR PROSITE; PS00473; GnRH; 1.  
FT NON\_TER 1 64  
SQ SEQUENCE 64 AA; 7028 MW; 1F91FFFB2B4BB6D0 CRC64;

Query Match 32.1%; Score 52; DB 13; Length 64;  
Best Local Similarity 88.9%; Pred. No. 4.5;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 HWSYGLRPG 10  
Db ||||| | | | | |  
16 HWSYGLSPG 24

## RESULT 10

Q8JIF2 PRELIMINARY; PRT; 64 AA.  
AC Q8JIF2;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Gonadotropin-releasing hormone (Fragment).  
OS Pagrus major (Red sea bream) (Chrysophrys major).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;  
OC Sparidae; Pagrus.  
OX NCBI\_TaxID=143350;  
RN [1]  
RP SEQUENCE FROM N.A.

RA Kato M., Elmesiry G.E.;  
RT "Sequence comparison of GnRH genes in closely-related Sparidae fishes."  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB089314; BAC07230.1; -.  
DR InterPro; IPR002012; GnRH.  
DR InterPro; IPR004079; GonadolibrinI.  
DR Pfam; PF00446; GnRH; 1.  
DR PRINTS; PR01541; GONADOLIBRN1.  
DR PROSITE; PS00473; GnRH; 1.  
FT NON\_TER 1 64  
SQ SEQUENCE 64 AA; 6967 MW; 119F0D0CDD14BB6D0 CRC64;

Query Match 32.1%; Score 52; DB 13; Length 64;  
Best Local Similarity 88.9%; Pred. No. 4.5;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 HWSYGLRPG 10  
Db ||||| | | | | |  
16 HWSYGLSPG 24

## RESULT 11

Q8JIF4 PRELIMINARY; PRT; 68 AA.  
AC Q8JIF4;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Gonadotropin-releasing hormone (Fragment).  
OS Acanthopagrus latus (Yellowfin porgy).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;  
OC Sparidae; Acanthopagrus.  
OX NCBI\_TaxID=8177;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kato M., Elmesiry G.E.;  
RT "Sequence comparison of GnRH genes in closely-related Sparidae fishes."  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB089312; BAC07228.1; -.  
DR InterPro; IPR002012; GnRH.  
DR InterPro; IPR004079; GonadolibrinI.  
DR Pfam; PF00446; GnRH; 1.  
DR PRINTS; PR01341; GONADOLIBRN1.  
DR PROSITE; PS00473; GnRH; 1.  
FT NON\_TER 1 68  
SQ SEQUENCE 68 AA; 7543 MW; 067708609FE8E771 CRC64;

Query Match 32.1%; Score 52; DB 13; Length 68;  
Best Local Similarity 88.9%; Pred. No. 4.8;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 HWSYGLRPG 10  
Db ||||| | | | | |  
20 HWSYGLSPG 28

## RESULT 12

Q9YI26 PRELIMINARY; PRT; 87 AA.  
AC Q9YI26;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE SbgNHR (Gonadolibrin) (Gonadotropin-releasing hormone) (LH-RH) (Luliberin) (Fragment).  
DE Sparus aurata (Gilthead sea bream).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Query Match 32.4%; Score 52.5; DB 16; Length 214;

Actinop  
Anguill

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:29:11 ; Search time 30.0554 Seconds  
(without alignments)  
266.163 Million cell updates/sec

Title: US-09-848-834A-15  
Perfect score: 162  
Sequence: 1 XHWSYGLRPGSSGPSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriaph.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	76	46.9	1310	2 Q93N27	Q93n27 clostridium
2	59.5	36.7	90	13 Q90Y63	Q90y63 rana catesb
3	59	36.4	91	13 Q9PRH0	Q9prh0 anquilla ja
4	53.5	33.0	374	16 Q92YR6	Q92yr6 rhizobium m
5	53	32.7	358	17 Q8TZU8	Q8tzu8 pyrococcus
6	52.5	32.4	214	16 Q8E7H5	Q8e7h5 streptococc
7	52.5	32.4	214	16 Q8E2I9	Q8e2i9 streptococc
8	52.5	32.4	437	16 Q8XRN3	Q8xrn3 clostridium
9	52	32.1	64	13 Q8JIF3	Q8jif3 dentex dent
10	52	32.1	64	13 Q8JIF2	Q8jif2 pagrus majo
11	52	32.1	68	13 Q8JIF4	Q8jif4 acanthopagr
12	52	32.1	87	13 Q9YI26	Q9yi26 sparus aura
13	52	32.1	94	13 Q8JFY3	Q8jfy3 oreochromis
14	52	32.1	96	13 Q8UW80	Q8uw80 verasper mo
15	51	31.5	208	16 Q8YR98	Q8yr98 anabaena sp
16	51	31.5	1545	16 Q9RDQ1	Q9rdq1 streptomyce

17	50.5	31.2	1494	11 Q88902	Q88902 rattus norv
18	50	30.9	192	10 Q8LGA9	Q8lga9 arabidopsis
19	50	30.9	215	11 Q8C8T5	Q8c8t5 mu musculus
20	50	30.9	224	6 Q28848	Q28848 sminthopsis
21	50	30.9	575	17 Q8TSG3	Q8tsq3 methanosarc
22	50	30.9	1331	10 Q9SF03	Q9sf03 arabidopsis
23	49.5	30.6	54	13 Q90W09	Q90w09 oncorhynch
24	49.5	30.6	100	16 Q8G4A3	Q8g4a3 bifidobacte
25	49.5	30.6	256	16 P72950	P72950 synechocyst
26	49.5	30.6	1485	11 Q8CIW4	Q8ciw4 mu musculus
27	49	30.2	75	16 Q8UK61	Q8uk61 agrobacteri
28	49	30.2	271	5 P91045	P91045 caenorhabdi
29	49	30.2	301	17 Q8THS2	Q8ths2 methanosarc
30	49	30.2	402	10 Q8VYH6	Q8vyh6 arabidopsis
31	49	30.2	444	16 Q8FDB7	Q8fdb7 escherichia
32	49	30.2	481	2 Q50210	Q50210 melittangiu
33	49	30.2	532	5 Q96671	Q96671 drosophila
34	49	30.2	532	5 Q9VU53	Q9vu53 drosophila
35	49	30.2	574	17 Q8TRV2	Q8trv2 methanosarc
36	49	30.2	618	16 Q8P8W9	Q8p8w9 xanthomonas
37	49	30.2	619	16 Q8PKS0	Q8pks0 xanthomonas
38	49	30.2	1047	3 Q9P6E3	Q9p6e3 neurospora
39	48	29.6	91	13 Q8JIIQ7	Q8jiq7 oryzias lat
40	48	29.6	91	13 Q9DGC8	Q9dgc8 oryzias lat
41	48	29.6	104	11 Q9JL82	Q9j182 mu musculus
42	48	29.6	145	6 Q28568	Q28568 ovis aries
43	48	29.6	146	17 Q27851	Q27851 methanobact
44	48	29.6	238	11 Q8VC16	Q8vc16 mu musculus
45	48	29.6	306	10 Q9SDN3	Q9sdn3 prunus dulc

#### ALIGNMENTS

RESULT 1

Q93N27 ID Q93N27 PRELIMINARY; PRT; 1310 AA.  
AC Q93N27;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Tetanus toxin (Fragment).  
OS Clostridium tetani.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shumin Z., Dianliang L.;  
RT "Cloning and sequence analysis of tetanus toxin gene."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF389424; AAK72964.2; .  
DR InterPro; IPR000395; Bontoxilysin.  
DR InterPro; IPR001064; Crystallin.  
DR InterPro; IPR006025; Zn.Mtpeptdse.  
DR Pfam; PF01742; Peptidase\_M27; 1.  
DR PRINTS; PR00760; BONTOXILYSIN.  
DR PRODOM; PD001963; Bontoxilysin; 1.  
DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
FT NON\_TER 1  
FT NON\_TER 1310 1310  
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;

Query Match 46.9%; Score 76; DB 2; Length 1310;  
Best Local Similarity 93.8%; Pred. No. 0.039;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 16 LOYIKANSKFIGITEL 31  
:|||||  
DB 830 MOYIKANSKFIGITEL 845



QY 18 EKIKAKMEKASSVFNVNS 36  
Db 3 EKIKAKMEKASSVFNVNS 21

## RESULT 15

US-08-446-692-27  
; Sequence 27, Application US/08446692  
; Patent No. 5759551  
; GENERAL INFORMATION:  
; APPLICANT: Ladd, Anna  
; APPLICANT: Wang, Chang Yi  
; APPLICANT: Zamb, Timothy  
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maria C.H. Lin  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: US  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,692  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria C.H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)415-8745  
; TELEFAX: (516)751-6849  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 33 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-446-692-27

Query Match 48.6%; Score 89; DB 1; Length 33;  
Best Local Similarity 100.0%; Pred. No. 5.5e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKIKAKMEKASSVFNVNS 36  
Db 3 EKIKAKMEKASSVFNVNS 21

Search completed: September 24, 2003, 17:00:24  
Job time : 10.7704 secs

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/827,682  
;; FILING DATE: 29-JAN-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/749,568  
;; FILING DATE: 26-AUG-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Parmelee, Steven W.  
;; REGISTRATION NUMBER: 31,990  
;; REFERENCE/DOCKET NUMBER: 14137-26-4  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206) 467-9600  
;; TELEFAX: (206) 623-6793  
;; INFORMATION FOR SEQ ID NO: 97:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 21 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: Peptide  
;; LOCATION: 1..21  
;; OTHER INFORMATION: /note= "Malaria circumsporozoite"  
;; OTHER INFORMATION: 378-398"  
US-08-197-484-97

Query Match 48.6% Score 89; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.3e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAMEKASSVFNVNVNS 36  
Db 3 EKKIAMEKASSVFNVNVNS 21

RESULT 13  
PCT-US95-02121-97  
;; Sequence 97, Application PC/TUS9502121  
;; GENERAL INFORMATION:  
;; APPLICANT:  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
;; TITLE OF INVENTION: CTL IMMUNITY  
;; NUMBER OF SEQUENCES: 153  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/02121  
;; FILING DATE: 16-FEB-1995  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/197,484  
;; FILING DATE: 16-FEB-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/935,811  
;; FILING DATE: 26-AUG-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/874,491  
;; FILING DATE: 27-APR-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/827,682  
;; FILING DATE: 29-JAN-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/749,568  
;; FILING DATE: 26-AUG-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Parmelee, Steven W.  
;; REGISTRATION NUMBER: 31,990  
;; REFERENCE/DOCKET NUMBER: 14137-26-4PC  
;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (206) 467-9600  
;; TELEFAX: (415) 543-5043  
;; INFORMATION FOR SEQ ID NO: 97:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 21 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: Peptide  
;; LOCATION: 1..21  
;; OTHER INFORMATION: /note= "Malaria circumsporozoite"  
;; OTHER INFORMATION: 378-398"  
PCT-US95-02121-97

Query Match 48.6% Score 89; DB 5; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.3e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAMEKASSVFNVNVNS 36  
Db 3 EKKIAMEKASSVFNVNVNS 21

RESULT 14  
PCT-US95-13841-20  
;; Sequence 20, Application PC/TUS9513841  
;; GENERAL INFORMATION:  
;; APPLICANT: United Biomedical Inc; Walfield, Alan M.;  
;; APPLICANT: Wang, Chang Yi  
;; TITLE OF INVENTION: Synthetic IgE Membrane Anchor  
;; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy  
;; NUMBER OF SEQUENCES: 32  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Maria C.H. Lin  
;; STREET: 345 Park Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10154  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: WordPerfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/13841  
;; FILING DATE: 25-OCT-1995  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/328,519  
;; FILING DATE: 25-OCT-1994  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Lin, Maria C.H.  
;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4117  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-758-4800  
;; TELEFAX: 212-751-6849  
;; TELEX: 421792  
;; INFORMATION FOR SEQ ID NO: 20:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 21 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
PCT-US95-13841-20

Query Match 48.6% Score 89; DB 5; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.3e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAMEKASSVFNVNVNS 36  
Db 3 EKKIAMEKASSVFNVNVNS 21

TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-100-409A-54

Query Match 48.6%; Score 89; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.3e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36  
Db 3 EKKIAKMEKASSVFNVNS 21  
|||||

RESULT 10  
US-08-464-496-17  
Sequence 17, Application US/08464496  
Patent No. 6322789

GENERAL INFORMATION:  
APPLICANT: Epiimmune, Inc.  
APPLICANT: Vitello, Maria  
APPLICANT: Chesnut, Robert  
TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL  
TITLE OF INVENTION: EPITOPES  
FILE REFERENCE: 39963-20001.13  
CURRENT APPLICATION NUMBER: US/08/464,496  
CURRENT FILING DATE: 1995-06-05  
PRIOR APPLICATION NUMBER: 07/935,811  
PRIOR FILING DATE: 1992-08-26  
PRIOR APPLICATION NUMBER: 07/874,491  
PRIOR FILING DATE: 1992-04-27  
PRIOR APPLICATION NUMBER: 07/827,682  
PRIOR FILING DATE: 1992-01-29  
PRIOR APPLICATION NUMBER: 07/749,568  
PRIOR FILING DATE: 1991-08-26  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17  
LENGTH: 21  
TYPE: PRT  
ORGANISM: Malaria circumsporozoite 378-398  
US-08-464-496-17

Query Match 48.6%; Score 89; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.3e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36  
Db 3 EKKIAKMEKASSVFNVNS 21  
|||||

RESULT 11  
US-08-788-822A-12  
Sequence 12, Application US/08788822A  
Patent No. 6413935

GENERAL INFORMATION:  
APPLICANT: Alexander, Jeffrey L.  
APPLICANT: DeFrees, Shawn  
APPLICANT: Sette, Alessandro  
TITLE OF INVENTION: Induction of Immune Response Against  
TITLE OF INVENTION: Desired Determinants  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA

ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/788,822A  
FILING DATE: 23-JAN-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/010,510  
FILING DATE: 24-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 014137-009210US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-788-822A-12

Query Match 48.6%; Score 89; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.3e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36  
Db 3 EKKIAKMEKASSVFNVNS 21  
|||||

RESULT 12  
US-08-197-484-97  
Sequence 97, Application US/08197484  
Patent No. 6419931  
GENERAL INFORMATION:  
APPLICANT: VITELLO, Maria A.  
APPLICANT: CHESTNUT, Robert W.  
APPLICANT: SETTE, Alessandro D.  
APPLICANT: CELIS, Esteban  
APPLICANT: GRAY, Howard  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
TITLE OF INVENTION: CTL IMMUNITY  
NUMBER OF SEQUENCES: 153  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Steuart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/197,484  
FILING DATE: 16-FEB-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/935,811  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992

```
US-08-446-692-48
; Sequence 48, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-48

Query Match 48.6%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKIAKMEKASSVFNVNS 36
Db 3 EKIAKMEKASSVFNVNS 21

RESULT 8
US-08-488-351A-48
; Sequence 48, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

US-09-848-834a-16.rai
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-48

Query Match 48.6%; Score 89; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKIAKMEKASSVFNVNS 36
Db 3 EKIAKMEKASSVFNVNS 21

RESULT 9
US-09-100-409A-54
; Sequence 54, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
```



## RESULT 2

US-08-760-797A-1  
; Sequence 1, Application US/08760797A  
; Patent No. 5928902  
; GENERAL INFORMATION:  
; APPLICANT: De Wilde, Michel  
; TITLE OF INVENTION: Hybrid Protein Between CS  
; TITLE OF INVENTION: from Plasmodium and HbsAG  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Smithkline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/760,797A  
; FILING DATE: 04-DEC-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/442,612  
; FILING DATE: 17-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: B45015-1C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 423 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-760-797A-1

Query Match 49.7%; Score 91; DB 2; Length 423;  
Best Local Similarity 55.3%; Pred. No. 5.2e-05;  
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRPSSGPSLD-----EKKIAMEKASSVFNVNVS 36  
:||||: | |||| |||| |||| |||| ||||  
Db 155 IKPGSANKPKDLDYNDIEKKICKMEKCSVFNVNVS 192

## RESULT 3

US-08-760-797A-3  
; Sequence 3, Application US/08760797A  
; Patent No. 5928902  
; GENERAL INFORMATION:  
; APPLICANT: De Wilde, Michel  
; TITLE OF INVENTION: Hybrid Protein Between CS  
; TITLE OF INVENTION: from Plasmodium and HbsAG  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Smithkline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/760,797A  
; FILING DATE: 04-DEC-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/442,612  
; FILING DATE: 17-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: B45015-1C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 424 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-760-797A-3

Query Match 49.7%; Score 91; DB 2; Length 424;  
Best Local Similarity 55.3%; Pred. No. 5.3e-05;  
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRPSSGPSLD-----EKKIAMEKASSVFNVNVS 36  
:||||: | |||| |||| |||| |||| ||||  
Db 152 IKPGSANKPKDLDYNDIEKKICKMEKCSVFNVNVS 189

## RESULT 4

US-08-932-929B-1  
; Sequence 1, Application US/08932929B  
; Patent No. 6169171  
; GENERAL INFORMATION:  
; APPLICANT: De Wilde, Michel  
; APPLICANT: Cohen, Joseph  
; TITLE OF INVENTION: Hybrid Protein Between CS  
; TITLE OF INVENTION: from Plasmodium and HbsAG  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Smithkline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/932,929B  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/760,797  
; FILING DATE: 04-DEC-1996  
; APPLICATION NUMBER: 08/442,612  
; FILING DATE: 17-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Baumeister, Kirk  
; REGISTRATION NUMBER: 33,833  
; REFERENCE/DOCKET NUMBER: B45015-1FWC2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5096

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:46:20 ; Search time 10.6704 Seconds  
(without alignments)  
142.750 Million cell updates/sec

Title: US-09-848-834A-16  
Perfect score: 183  
Sequence: 1 XHWSYGLRPGSSGSLDEKTKIAKMEKASSVFNVNS 36

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues  
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pap:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description
1	91	49.7	412	1	US-08-313-288B-18
2	91	49.7	423	2	US-08-760-797A-1
3	91	49.7	424	2	US-08-760-797A-3
4	91	49.7	424	3	US-08-932-929B-1
5	91	49.7	424	3	US-08-932-929B-3
6	89	48.6	21	1	US-08-186-266-6
7	89	48.6	21	1	US-08-446-692-48
8	89	48.6	21	3	US-08-488-351A-48
9	89	48.6	21	3	US-09-100-409A-54
10	89	48.6	21	4	US-08-464-496-17
11	89	48.6	21	4	US-08-788-822A-12
12	89	48.6	21	4	US-08-197-484-97
13	89	48.6	21	5	PCT-US95-02121-97
14	89	48.6	21	5	PCT-US95-13841-20
15	89	48.6	33	1	US-08-446-692-27
16	89	48.6	33	2	US-08-488-351A-27
17	79	43.2	17	4	US-08-464-496-16
18	79	43.2	17	4	US-08-197-484-96
19	79	43.2	17	5	PCT-US95-02121-96
20	75	41.0	16	2	US-08-817-933A-7
21	73.5	40.2	20	1	US-08-465-167A-20
22	73.5	40.2	20	4	US-08-627-820-20
23	73.5	40.2	20	5	PCT-US92-07218-17
24	71	38.8	17	1	US-08-188-223-6
25	71	38.8	17	3	US-08-968-466-6
26	71	38.8	17	4	US-08-478-546B-6
27	70	38.3	15	6	5169933-30

28	64	35.0	699	2	US-08-694-865-16	Sequence 16, Appl
29	64	35.0	699	3	US-09-124-491-16	Sequence 16, Appl
30	64	35.0	699	4	US-09-383-912-16	Sequence 16, Appl
31	63.5	34.7	16	1	US-08-465-167A-19	Sequence 19, Appl
32	63.5	34.7	16	4	US-08-627-820-19	Sequence 19, Appl
33	63.5	34.7	16	5	PCT-US92-07218-16	Sequence 16, Appl
34	62	33.9	14	1	US-08-453-588-23	Sequence 23, Appl
35	62	33.9	14	3	US-08-521-079-23	Sequence 23, Appl
36	62	33.9	17	1	US-07-690-983D-18	Sequence 18, Appl
37	62	33.9	18	1	US-07-690-983D-28	Sequence 28, Appl
38	62	33.9	21	1	US-08-305-871A-12	Sequence 12, Appl
39	62	33.9	49	1	US-08-387-156-4	Sequence 4, Appl
40	62	33.9	49	2	US-08-694-865-4	Sequence 4, Appl
41	62	33.9	49	3	US-08-878-748-4	Sequence 4, Appl
42	62	33.9	49	3	US-09-124-491-4	Sequence 4, Appl
43	62	33.9	49	4	US-09-383-912-4	Sequence 4, Appl
44	62	33.9	544	1	US-08-387-156-10	Sequence 10, Appl
45	62	33.9	544	2	US-08-694-865-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-08-313-288B-18  
; Sequence 18, Application US/08313288B  
; Patent No. 5750502  
; GENERAL INFORMATION:  
; APPLICANT: Jessell, Thomas M. and Avi Hu Klar  
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A  
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/313,288B  
; FILING DATE: January 5, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0526  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 412 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-313-288B-18

Query Match 49.7%; Score 91; DB 1; Length 412;  
Best Local Similarity 55.3%; Pred. NO. 5.1e-05;  
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;  
QY 7 LRPSSGSLD-----EKKIAKMEKASSVFNVNS 36  
:::|||||  
DB 361 IKPGSANKPKDLDYENDIERKICKMEKCSSVFNVNS 398

; LOCATION: (46)..(46)  
; OTHER INFORMATION: Amidated glycine or glycynamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(31)  
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor  
; NAME/KEY: PEPTIDE  
; LOCATION: (32)..(37)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (38)..(46)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-19

Query Match 48.6%; Score 89; DB 9; Length 46;  
Best Local Similarity 63.3%; Pred. No. 1.7e-05;  
Matches 19; Conservative 2; Mismatches 3; Indels 6; Gaps 1;

Qy 2 HWSYGLRPGSGPSLDEKKIAKMEKASSVF 31  
|||||  
Db 2 HWSYGLRPGSGPSL-----QYIKANSKF 25

Search completed: September 24, 2003, 17:33:09  
Job time : 18.0499 secs

RESULT 13  
US-10-128-711-97  
; Sequence 97, Application US/10128711  
; Publication No. US20030099634A1  
; GENERAL INFORMATION:  
; APPLICANT: VITIELLO, Maria A.  
; CHESTNUT, Robert W.  
; SETTE, Alessandro D.  
; CELIS, Esteban  
; GRAY, Howard  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Townsend and Townsend Kourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/128,711  
; FILING DATE: 22-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/197,484  
; FILING DATE: 16-FEB-1994  
; APPLICATION NUMBER: US 07/935,811  
; FILING DATE: 26-AUG-1992  
; APPLICATION NUMBER: US 07/874,491  
; FILING DATE: 27-APR-1992  
; APPLICATION NUMBER: US 07/827,682  
; FILING DATE: 29-JAN-1992  
; APPLICATION NUMBER: US 07/749,568  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; REFERENCE/DOCKET NUMBER: 14137-26-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 467-9600  
; TELEFAX: (206) 623-6793  
; INFORMATION FOR SEQ ID NO: 97:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..21  
; OTHER INFORMATION: /note= "Malaria circumsporozoite  
; 378-398"  
; SEQUENCE DESCRIPTION: SEQ ID NO: 97:  
US-10-128-711-97

Query Match 48.6%; Score 89; DB 15; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.7e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 18 EKKIAKMEKASSVFNVVNS 36  
Db 3 EKKIAKMEKASSVFNVVNS 21

RESULT 14  
US-09-848-834A-15  
; Sequence 15, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15  
; LENGTH: 31  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the  
; OTHER INFORMATION: RH hormone linked by a spacer to amino sequence 830-844 of the  
; OTHER INFORMATION: tanus toxoid precursor (Tentoxylisin)  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(31)  
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor  
; OTHER INFORMATION: (Tentoxylisin)  
US-09-848-834A-15  
Query Match 48.6%; Score 89; DB 9; Length 31;  
Best Local Similarity 63.3%; Pred. No. 1.1e-05;  
Matches 19; Conservative 2; Mismatches 3; Indels 6; Gaps 1;  
QY 2 HWSYGLRPGSSGSLDKKIAKMEKASSVF 31  
Db 2 HWSYGLRPGSSGSL-----QYIKANSKF 25  
RESULT 15  
US-09-848-834A-19  
; Sequence 19, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19  
; LENGTH: 46  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human  
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of the  
; OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino  
; OTHER INFORMATION: sequence 1-10 of GnRH  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: MOD\_RES

Qy 18 EKKIAMEKASSVFNVNVS 36  
| | | | | | | | | |  
db 3 EKKIAMEKASSVFNVNVS 21

## RESULT 5

US-09-820-843A-31  
; Sequence 31, Application US/09820843A  
; Publication No. US20030039963A1  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN  
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES  
; FILE REFERENCE: Q63915  
; CURRENT APPLICATION NUMBER: US/09/820,843A  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 31  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Circumsporozoite (CS) protein  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: gi|4493889  
US-09-820-843A-31

Query Match 49.7%; Score 91; DB 11; Length 396;

Best Local Similarity 55.3%; Pred. No. 0.00011; Indels 8; Gaps 1;

Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSLD-----EKKIAKMEKASSVFNVNS 36

Db 345 IKPGSANKPKDLDYANDIEKKICKMEKCSSVFNVNS 382

## RESULT 6

US-10-239-313A-54  
; Sequence 54, Application US/10239313A  
; Publication No. US20030175285A1  
; GENERAL INFORMATION:  
; APPLICANT: KLINGUER - HAMOUR, Christine  
; APPLICANT: CORVAIA, Nathalie  
; APPLICANT: BECK, Alain  
; APPLICANT: GOETSCH, Lilliane  
; TITLE OF INVENTION: MOLECULE OF PHARMACEUTICAL INTEREST COMPRISING AT ITS  
; TITLE OF INVENTION: N-TERMINAL A GLUTAMIC ACID OR A GLUTAMINE IN THE FORM  
; TITLE OF INVENTION: OF A PHYSIOLOGICALLY ACCEPTABLE STRONG ACID  
; FILE REFERENCE: 343 727 - US  
; CURRENT APPLICATION NUMBER: US/10/239,313A  
; CURRENT FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: FR 00/03711  
; PRIOR FILING DATE: 2000-03-23  
; PRIOR APPLICATION NUMBER: PCT 01/70772  
; PRIOR FILING DATE: 2001-03-22  
; NUMBER OF SEQ ID NOS: 697  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 54  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Plasmodium malariae  
US-10-239-313A-54

Query Match 48.6%; Score 89; DB 12; Length 19;

Best Local Similarity 100.0%; Pred. No. 6e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36

Db 1 EKKIAKMEKASSVFNVNS 19

## RESULT 7

US-09-932-165-1482

; Sequence 1482, Application US/09932165  
; Publication No. US20030134784A1  
; GENERAL INFORMATION:

; APPLICANT: RAITANO, ARTHUR

; APPLICANT: CHALLITA-EID, PIA M.

; APPLICANT: FARIS, MARY

; APPLICANT: SAFRAN, DOUGLAS

; APPLICANT: AFAR, DANIEL

; APPLICANT: LEVIN, ELANA

; APPLICANT: HUBERT, RENE

; APPLICANT: GE, WANGMAO

; TITLE OF INVENTION: JAKOBOVITS, AYA

; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
; TITLE OF INVENTION: 83P2H3 AND CatF2E11 USEFUL IN TREATMENT AND

; TITLE OF INVENTION: DETECTION OF CANCER

; FILE REFERENCE: 51158-20014.00

; CURRENT APPLICATION NUMBER: US/09/932,165

; CURRENT FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: 60/226,329

; PRIOR FILING DATE: 2000-08-17

; NUMBER OF SEQ ID NOS: 1508

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1482

; LENGTH: 21

; TYPE: PRT

; ORGANISM: Plasmodium falciparum

US-09-932-165-1482

Query Match 48.6%; Score 89; DB 12; Length 21;

Best Local Similarity 100.0%; Pred. No. 6.7e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36

Db 3 EKKIAKMEKASSVFNVNS 21

## RESULT 8

US-10-116-118-33  
; Sequence 33, Application US/10116118  
; Publication No. US20030143672A1  
; GENERAL INFORMATION:  
; APPLICANT: Tangri, Shabnam  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn  
; APPLICANT: Fikes, John D.  
; TITLE OF INVENTION: Heteroclitic Analogs and Related Methods  
; FILE REFERENCE: 2060.0090003  
; CURRENT APPLICATION NUMBER: US/10/116,118  
; CURRENT FILING DATE: 2002-08-07  
; PRIOR APPLICATION NUMBER: US 60/166,529  
; PRIOR FILING DATE: 1999-11-18  
; PRIOR APPLICATION NUMBER: US 60/239,008  
; PRIOR FILING DATE: 2000-10-06  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 33  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-10-116-118-33

Query Match 48.6%; Score 89; DB 12; Length 21;

Best Local Similarity 100.0%; Pred. No. 6.7e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36

Db 3 EKKIAKMEKASSVFNVNS 21

## RESULT 9

US-09-935-384-710



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:57:10 ; Search time 18.0499 Seconds  
(without alignments)  
301.778 Million cell updates/sec

Title: US-09-848-834A-16

Perfect score: 183

Sequence: 1 XHWSYGLRPGSGPSLDEKKIARMEKASSVFNVNS 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	182	99.5	36	US-09-848-834A-16	Sequence 16, Appl
2	182	99.5	51	US-09-848-834A-20	Sequence 20, Appl
3	95	51.9	20	US-09-848-834A-3	Sequence 3, Appl
4	95	51.9	33	US-09-848-834A-12	Sequence 12, Appl
5	91	49.7	396	US-09-820-843A-31	Sequence 31, Appl
6	89	48.6	19	US-10-239-313A-54	Sequence 54, Appl
7	89	48.6	21	US-09-932-165-1482	Sequence 1482, Ap
8	89	48.6	21	US-10-116-118-33	Sequence 33, Appl
9	89	48.6	21	US-09-935-384-710	Sequence 710, App
10	89	48.6	21	US-10-062-109A-761	Sequence 761, App
11	89	48.6	21	US-09-942-052-711	Sequence 711, App
12	89	48.6	21	US-10-001-469-1404	Sequence 1404, Ap
13	89	48.6	21	US-10-128-711-97	Sequence 97, Appl
14	89	48.6	31	US-09-848-834A-15	Sequence 15, Appl
15	89	48.6	46	US-09-848-834A-19	Sequence 19, Appl

16	87	47.5	34	9	US-09-848-834A-13	Sequence 13, Appl
17	87	47.5	37	9	US-09-848-834A-14	Sequence 14, Appl
18	87	47.5	47	9	US-09-848-834A-17	Sequence 17, Appl
19	87	47.5	50	9	US-09-848-834A-18	Sequence 18, Appl
20	82	44.8	21	11	US-09-747-802-23	Sequence 23, Appl
21	82	44.8	21	11	US-09-865-294-15	Sequence 15, Appl
22	79	43.2	17	12	US-10-239-313A-53	Sequence 53, Appl
23	79	43.2	17	15	US-10-128-711-96	Sequence 96, Appl
24	75	41.0	17	12	US-10-239-313A-36	Sequence 36, Appl
25	75	41.0	17	12	US-10-239-313A-38	Sequence 38, Appl
26	75	41.0	17	12	US-10-239-313A-39	Sequence 39, Appl
27	75	41.0	17	12	US-10-239-313A-41	Sequence 41, Appl
28	75	41.0	17	12	US-10-239-313A-60	Sequence 60, Appl
29	75	41.0	17	12	US-10-239-313A-64	Sequence 64, Appl
30	75	41.0	17	12	US-10-239-313A-67	Sequence 67, Appl
31	74	40.4	17	12	US-10-239-313A-59	Sequence 59, Appl
32	74	40.4	17	12	US-10-239-313A-65	Sequence 65, Appl
33	74	40.4	17	12	US-10-239-313A-66	Sequence 66, Appl
34	73	39.9	17	12	US-10-239-313A-14	Sequence 14, Appl
35	73	39.9	17	12	US-10-239-313A-26	Sequence 26, Appl
36	73	39.9	17	12	US-10-239-313A-31	Sequence 31, Appl
37	73	39.9	17	12	US-10-239-313A-32	Sequence 32, Appl
38	73	39.9	17	12	US-10-239-313A-37	Sequence 37, Appl
39	73	39.9	17	12	US-10-239-313A-40	Sequence 40, Appl
40	73	39.9	17	12	US-10-239-313A-44	Sequence 44, Appl
41	73	39.9	17	12	US-10-239-313A-47	Sequence 47, Appl
42	73	39.9	17	12	US-10-239-313A-49	Sequence 49, Appl
43	73	39.9	17	12	US-10-239-313A-52	Sequence 52, Appl
44	73	39.9	17	12	US-10-239-313A-58	Sequence 58, Appl
45	73	39.9	17	12	US-10-239-313A-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1

US-09-848-834A-16  
; Sequence 16, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the  
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 378-398 of  
; OTHER INFORMATION: the Plasmodium falciparum circumsporozoite (CSP) protein  
; NAME/KEY: MOD\_RES:  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(36)  
; OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria  
; OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite  
; OTHER INFORMATION: (CSP) protein  
US-09-848-834A-16



```

ID Q819J1 PRELIMINARY; PRT: 393 AA.
AC Q819J1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Yaounde;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI. Asebo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL: AF540444; AAN87578.1;
SQ SEQUENCE 393 AA; 42276 MW; 7D2C7CCFAA3153A9 CRC64;

Query Match 51.4%; Score 94; DB 5; Length 393;
Best Local Similarity 60.5%; Pred. No. 5.3e-05;
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 2;

QY 7 LRPSSGP---SLD-----EKKIAMEKASSVFNVVNS 36
DB 342 IKPGSAGKPNELDYNDIEKKICKMEKCSSVFNVVNS 379

RESULT 15
Q8MZK7
ID Q8MZK7 PRELIMINARY; PRT: 69 AA.
AC Q8MZK7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Circumsporozoite protein II (Fragment).
GN CSPII.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FCCL/HN;
RA Chen H.H., Yu X.B., Wu Z.D., Xu J.;
RT "Gene cloning of circumsporozoite protein (CSP) II gene from
RT Plasmodium falciparum (FCCL/HN).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY093672; AAM19072.1;
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
DR PROSITE; PS00092; TSPI; 1.
FT NON_TER 1
SQ SEQUENCE 69 AA; 7670 MW; F3E1AA54671FE6FF CRC64;

Query Match 49.7%; Score 91; DB 5; Length 69;
Best Local Similarity 55.3%; Pred. No. 2.1e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGPSLD-----EKKIAMEKASSVFNVVNS 36
DB 18 IKPGSANKPRDELNDYNDIEKKICKMEKCSSVFNVVNS 55

```

Search completed: September 24, 2003, 17:47:21  
Job time : 35.903 secs

[illegible]

QV 7 LRPSSGPSLD-----EKKIAKMEKASSVFNVNS 36

```
ID Q25729 PRELIMINARY; PRT; 408 AA.
AC Q25729;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Circumsporozoite protein.
CS Plasmodium falciparum.
OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Santa Lucia;
RA Qari S.H., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20969; AAA63153.1; -
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria.
SQ SEQUENCE 408 AA; 43871 MW; A545BE517822515C CRC64;

Query Match 55.7%; Score 102; DB 5; Length 408;
Best Local Similarity 60.5%; Pred. No. 3.9e-06;
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

QY 7 LRPSSGPSLD-----EKKIAMEKASSVFNVNVS 36
DB 357 IKPGSAGSKDELVDYNDIEKKICKMEKCSVFNVNVS 394

RESULT 3
Q27325
ID Q27325 PRELIMINARY; PRT; 436 AA.
AC Q27325;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Circumsporozoite protein.
CS Plasmodium falciparum.
OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RA Dane J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RA la Cruz V.F.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RA Jongwutives S., Tanabe K., Kanbara H.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M83164; AAA29542.1; -
DR EMBL; M83150; AAA29563.1; -
DR EMBL; M83163; AAA29576.1; -
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
```

```
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 436 AA; 46688 MW; 5B42FF3348B68655 CRC64;

Query Match 55.7%; Score 102; DB 5; Length 436;
Best Local Similarity 60.5%; Pred. No. 4.2e-06;
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

QY 7 LRPSSGPSLD-----EKKIAMEKASSVFNVNVS 36
DB 385 IKPGSAGSKDELVDYNDIEKKICKMEKCSVFNVNVS 422

RESULT 4
Q9U0Q2
ID Q9U0Q2 PRELIMINARY; PRT; 79 AA.
AC Q9U0Q2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Circumsporozoite protein (fragment).
CS Plasmodium falciparum.
OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D51;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269945; CAB64243.1; -
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1
FT NON_TER 79
SQ SEQUENCE 79 AA; 8893 MW; 56FBFA76D859B416 CRC64;

Query Match 53.6%; Score 98; DB 5; Length 79;
Best Local Similarity 59.5%; Pred. No. 2.4e-06;
Matches 22; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

QY 7 LRPSSGPSLD-----EKKIAMEKASSVFNVNVS 35
DB 43 IKPGSAGSKDELVDYNDIEKKICKMEKCSVFNVNVS 79

RESULT 5
Q8I9H8
ID Q8I9H8 PRELIMINARY; PRT; 392 AA.
AC Q8I9H8;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Circumsporozoite protein.
CS Plasmodium falciparum.
OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Asembo Bay;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
```

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	102	55.7	117	5	Q25797	Q25797	plasmidium
2	102	55.7	408	5	Q25729	Q25729	plasmidium
3	102	55.7	436	5	Q27325	Q27325	plasmidium
4	98	53.6	79	5	Q9U0Q2	Q9U0Q2	plasmidium
5	98	53.6	392	5	Q819H8	Q819H8	plasmidium
6	98	53.6	396	5	Q819I1	Q819I1	plasmidium
7	98	53.6	396	5	Q819I0	Q819I0	plasmidium
8	98	53.6	396	5	Q819H9	Q819H9	plasmidium
9	97	53.0	360	5	Q819H7	Q819H7	plasmidium
10	97	53.0	360	5	Q819H6	Q819H6	plasmidium
11	97	53.0	412	5	Q819H5	Q819H5	plasmidium
12	97	53.0	420	5	Q25838	Q25838	plasmidium
13	94	51.4	389	5	Q819J2	Q819J2	plasmidium
14	94	51.4	393	5	Q819J1	Q819J1	plasmidium
15	91	49.7	69	5	Q8M2K7	Q8M2K7	plasmidium
16	91	49.7	80	5	Q9U0P3	Q9U0P3	plasmidium



Best Local Similarity 37.1%; Pred. No. 0.99;  
Matches 13; Conservative 9; Mismatches 7; Indels 6; Gaps 1;

```
db 291 RKGSNKKAEADLTLEDIDTEICKMDKCSSIFNIVSN 325
```

Job time : 6.78144 secs



RA Tan L., Rousseau P.;  
RT "The chemical identity of the immunoreactive LHRH-like peptide  
RL biosynthesized in the human placenta.";  
RN Biochem. Biophys. Res. Commun. 109:1061-1071(1982).  
[5]  
RP VARIANT SER-16.  
RX MEDLINE=99318093; PubMed=10391209;  
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
RA Shaw N., Lane C.R., Lim E.P., Kalyanarajan N., Nemesh J., Ziaugra L.,  
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
RA Lander E.S.;  
RT "Characterization of single-nucleotide polymorphisms in coding regions  
of human genes.";  
RL Nat. Genet. 22:231-238(1999).  
[6]  
RN ERRATUM.  
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
RA Shaw N., Lane C.R., Lim E.P., Kalyanarajan N., Nemesh J., Ziaugra L.,  
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
RA Lander E.S.;  
RL Nat. Genet. 23:373-373(1999).  
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- PHARMACEUTICAL: Available under the names Factrel (Averis Labs),  
CC Lutepulse or Lutrelief (Ferring Pharmaceuticals) and Relisorm  
CC (Serono).  
CC -!- SIMILARITY: Belongs to the GNRH family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; X01059; CAA25526.1; -;  
CC EMBL; M12578; AAA35916.1; -;  
CC EMBL; X15215; CAA33285.1; -;  
CC PIR; S05308; RHUHG.  
CC Genew; HGNC:4419; GNRH1.  
CC MIM; 152760; -;  
CC GO; GO:0005625; C:soluble fraction; TAS.  
CC GO; GO:0005183; P:luteinizing hormone-releasing factor activity; TAS.  
CC GO; GO:0007267; P:cell-cell signaling; TAS.  
CC GO; GO:0007275; P:development; TAS.  
CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
CC GO; GO:0007165; P:signal transduction; TAS.  
CC InterPro; IPR002012; GNRH.  
CC InterPro; IPR004079; GonadoliberinI.  
CC Pfam; PF00446; GNRH; 1.  
CC PRINTS; PR01541; GONADOLIBERNI.  
CC PROSITE; PS00473; GNRH; 1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Placenta; Pharmaceutical; Signal; Polymorphism;  
KW Pyrrolidone carboxylic acid.  
FT SIGNAL 1 23  
FT CHAIN 24 92  
FT PEPTIDE 24 33  
FT PEPTIDE 37 92  
FT ACT\_SITE 26 26  
FT  
FT MOD\_RES 24 24  
FT MOD\_RES 33 33  
FT VARIANT 16 16  
FT  
FT SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;  
Query Match 31.7%; Score 58; DB 1; Length 92;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
DB 25 HWSYGLRPG 33  
  
RESULT 13  
GONL\_RAT  
ID GONL\_RAT STANDARD; PRT; 92 AA.  
AC P07490;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
DE hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor  
DE I].  
GN GNRH1 OR GNRH.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86094338; PubMed=2867548;  
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;  
RT "Isolation of the gene and hypothalamic cDNA for the common precursor  
RT of gonadotropin-releasing hormone and prolactin release-inhibiting  
RT factor in human and rat.";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89384661; PubMed=2476669;  
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;  
RT "The rat gonadotropin-releasing hormone: SH locus: structure and  
RT hypothalamic expression.";  
RL Mol. Endocrinol. 3:1257-1262(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Thymus;  
RX MEDLINE=93105480; PubMed=1468115;  
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;  
RT "Thymocytes express a mRNA that is identical to hypothalamic  
RT luteinizing hormone-releasing hormone mRNA.";  
RL Cell. Mol. Neurobiol. 12:447-454(1992).  
RN [4]  
RP SEQUENCE OF 1-47 FROM N.A.  
RC TISSUE=Heart;  
RX MEDLINE=87149087; PubMed=3547652;  
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;  
RT "Two mammalian genes transcribed from opposite strands of the same  
RT DNA locus.";  
RL Science 235:1514-1517(1987).  
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.  
CC -!- SIMILARITY: Belongs to the GNRH family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; S08070; AAB24572.1; -;  
CC EMBL; M12579; AAA41263.1; -;  
CC EMBL; M31670; AAA41264.1; -;  
CC EMBL; M15527; AAA42141.1; ALT\_SEQ.  
CC EMBL; M15529; AAA42139.1; -;

DR PRINTS: PR01541; GONADOLIBRIN I.  
 DR PROSITE; PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Signal; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 21  
 FT CHAIN 22 90 PROGONADOLIBERIN I.  
 FT PEPTIDE 22 31 GONADOLIBERIN I.  
 FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.  
 FT ACT\_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.  
 FT MOD\_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA482654D9 CRC64;

Query Match 31.7%; Score 58; DB 1; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
 DB 23 HWSYGLRPG 31  
 |||||

RESULT 11  
 GONL\_PIG STANDARD; PRT; 91 AA.  
 AC P49921;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonaoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I].  
 GN GNRH1 OR GNRH.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RA Weesner G.D., Matteri R.L., Becker B.A.;  
 RL Submitted (MAY-1994) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 24-33.  
 RX MEDLINE=72114303; PubMed=4946067;  
 RA Baba Y., Matsuo H., Schally A.V.;  
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.  
 RT Confirmation of the proposed structure by conventional sequential  
 RT analyses.";  
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).  
 RN [3]  
 RP SYNTHESIS OF GONADOLIBERIN.  
 RX MEDLINE=72065376; PubMed=4942726;  
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;  
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-  
 RT phase method.";  
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).  
 RN [4]  
 RP SYNTHESIS OF GONADOLIBERIN.  
 RX MEDLINE=72117544; PubMed=4946275;  
 RA Baba Y., Arimura A., Schally A.V.;  
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";  
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the GNRH family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L32864; AAA31066.1; -  
 DR InterPro; IPR002012; GNRH.  
 DR InterPro; IPR004079; GonadoliberinI.  
 DR Pfam; PF00446; GNRH; 1.  
 DR PRINTS; PR01541; GONADOLIBRIN I.  
 DR PROSITE; PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Signal; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 23  
 FT CHAIN 24 91 PROGONADOLIBERIN I.  
 FT PEPTIDE 24 33 GONADOLIBERIN I.  
 FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.  
 FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.  
 FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;

Query Match 31.7%; Score 58; DB 1; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
 DB 25 HWSYGLRPG 33  
 |||||

RESULT 12  
 GONL\_HUMAN STANDARD; PRT; 92 AA.  
 AC P01148;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonaoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GNRH I) (Luliberin I) (Gonadorelin); GNRH-associated  
 DE peptide I].  
 GN GNRH1 OR GNRH OR LHRH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RA Haylick J.S., Adelman J.P., Seeburg P.H.;  
 RL "The complete nucleotide sequence of the human gonadotropin-releasing  
 RL hormone gene.";  
 RN Nucleic Acids Res. 17:6403-6403(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86094338; PubMed=2867548;  
 RA Adelman J.P., Mason A.J., Haylick J.S., Seeburg P.H.;  
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor  
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting  
 RT factor in human and rat.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANT SER-16.  
 RX MEDLINE=85012739; PubMed=6090951;  
 RA Seeburg P.H., Adelman J.P.;  
 RT "Characterization of cDNA for precursor of human luteinizing hormone  
 RT releasing hormone.";  
 RL Nature 311:666-668(1984).  
 RN [4]  
 RP SEQUENCE OF 24-33.  
 RX MEDLINE=83126573; PubMed=6760865;

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; S75918; AAB33096.1; -  
 PIR; I78541; I78541.  
 InterPro; IPR002012; GnRH.  
 InterPro; IPR004079; GonadoliberinI.  
 Pfam; PF00446; GnRH; 1.  
 PRINTS; PR01541; GONADOLIBERNI.  
 PROSITE; PS00473; GnRH; 1.  
 Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 Signal; Pyrrolidone carboxylic acid.  
 NON\_TER 1 1  
 SIGNAL <1 5 BY SIMILARITY.  
 CHAIN 6 >67 PROGONADOLIBERIN I.  
 PEPTIDE 6 15 GONADOLIBERIN I.  
 PEPTIDE 19 >67 GnRH-ASSOCIATED PEPTIDE I.  
 ACT\_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 ACTIVITY (BY SIMILARITY).  
 MOD\_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY  
 SIMILARITY).  
 MOD\_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY  
 SIMILARITY).  
 NON\_TER 67 67  
 SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 31.7%; Score 58; DB 1; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 0.16;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
 Db 7 HWSYGLRPG 15  
 |||||

## RESULT 9

GONL\_XENLA STANDARD; PRT; 89 AA.  
 ID GONL\_XENLA  
 AC P45656;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)  
 DE (LH-RH) (Luliberin I).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Forebrain;  
 RX MEDLINE=94185563; PubMed=8137750;  
 RA Hayes W.P., Wray S., Battey J.F.;  
 RT "The frog gonadotropin-releasing hormone-I (GnRH-I) gene has a  
 mammalian-like expression pattern and conserved domains in  
 GnRH-associated peptide, but brain onset is delayed until  
 metamorphosis.";  
 RL Endocrinology 134:1835-1844(1994).  
 CC !- FUNCTION: Stimulates the secretion of gonadotropins.  
 CC !- SUBCELLULAR LOCATION: Secreted.  
 CC !- SIMILARITY: Belongs to the GnRH family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L28040; AAA49728.1; -  
 DR PIR; I51423; I51423.  
 DR InterPro; IPR002012; GnRH.  
 DR InterPro; IPR004079; GonadoliberinI.  
 DR Pfam; PF00446; GnRH; 1.  
 DR PRINTS; PR01541; GONADOLIBERNI.  
 DR PROSITE; PS00473; GnRH; 1.  
 Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 Signal; Pyrrolidone carboxylic acid.  
 NON\_TER 1 23  
 SIGNAL 1 23  
 CHAIN 24 89 PROGONADOLIBERIN I.  
 PEPTIDE 24 33 GONADOLIBERIN I.  
 PEPTIDE 37 89 GONADOTROPIN-RELEASING HORMONE ASSOCIATED  
 PEPTIDE  
 FT PEPTIDE 37 85 GnRH-ASSOCIATED PEPTIDE I (GAP).  
 FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 89 AA; 10246 MW; 6F4F36FBAE0D4284 CRC64;

Query Match 31.7%; Score 58; DB 1; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10  
 Db 25 HWSYGLRPG 33  
 |||||

## RESULT 10

GONL\_MOUSE STANDARD; PRT; 90 AA.  
 ID GONL\_MOUSE  
 AC P13562;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor  
 DE I].  
 GN GnRH1 OR GnRH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87069928; PubMed=3024317;  
 RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,  
 RA Phillips H.S., Nikolics K., Seeburg P.H.;  
 RT "A deletion truncating the gonadotropin-releasing hormone gene is  
 responsible for hypogonadism in the hpg mouse.";  
 RL Science 234:1366-1371(1986).  
 CC !- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC !- SUBCELLULAR LOCATION: Secreted.  
 CC !- SIMILARITY: Belongs to the GnRH family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M14872; AAA37717.1; -  
 DR PIR; A47578; RHWG.  
 DR MGD; MGI:95789; GnRH.  
 DR InterPro; IPR002012; GnRH.  
 DR InterPro; IPR004079; GonadoliberinI.  
 DR Pfam; PF00446; GnRH; 1.

RT "Primary structure of the ovine hypothalamic luteinizing hormone-releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass spectrometry-decapeptide-Edman degradation).";

RL proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).

CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the GnRH family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U02517; AAA03433.1; --

DR InterPro; IPR002012; GnRH.

DR Pfam; PF00446; GnRH; 1.

DR PRINTS; PR01541; GONADOLIBRN1.

DR PROSITE; PS00473; GnRH; 1.

KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Pyrrolidone carboxylic acid.

FT NON\_TER 1 1

FT CHAIN 1 >61 PROGONADOLIBERIN I.

FT PEPTIDE 1 10 GONADOLIBERIN I.

FT PEPTIDE 14 >61 GnRH-ASSOCIATED PEPTIDE I.

FT ACT\_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.

FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

FT MOD\_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).

FT NON\_TER 61 61

SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;

Query Match 33.6%; Score 61.5; DB 1; Length 61;

Best Local Similarity 42.9%; Pred. No. 0.044;

Matches 15; Conservative 3; Mismatches 2; Indels 15; Gaps 2;

QY 2 HWSYGLRPGSGPSLDEKFKIAKMEKASVFNVYNS 36

Db 2 HWSYGLRPGG-----KRNK-----NVIDS 21

RESULT 7

GONL\_MESAU STANDARD; PRT; 63 AA.

AC 009163;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I) (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I] (Fragment).

DE (Fragment).

GN GnRH1 OR GnRH OR LHRH.

OS Mesocricetus auratus (Golden hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus

OX NCBI\_Taxid=10036;

RN [1]

RP SEQUENCE FROM N.A.

RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;

RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the GnRH family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U91938; AAB51302.1; --

DR InterPro; IPR002012; GnRH.

DR Pfam; PF00446; GnRH; 1.

DR PRINTS; PR01541; GONADOLIBRN1.

DR PROSITE; PS00473; GnRH; 1.

KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Pyrrolidone carboxylic acid.

FT NON\_TER 1 1

FT CHAIN 1 >63 PROGONADOLIBERIN I.

FT PEPTIDE 1 10 GONADOLIBERIN I.

FT PEPTIDE 14 >63 GnRH-ASSOCIATED PEPTIDE I (BY SIMILARITY).

FT ACT\_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY (BY SIMILARITY).

FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY).

FT MOD\_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY SIMILARITY).

FT NON\_TER 63 63

SQ SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;

Query Match 32.2%; Score 59; DB 1; Length 63;

Best Local Similarity 52.2%; Pred. No. 0.1;

Matches 12; Conservative 0; Mismatches 5; Indels 6; Gaps 1;

QY 2 HWSYGLRPGSS-----GPSLDE 18

Db 2 HWSYGLRPGGKRNARLGDSPQE 24

RESULT 8

GONL\_MACMU STANDARD; PRT; 67 AA.

AC P55247;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I) (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I] (Fragment).

DE (Fragment).

GN GnRH1 OR GnRH OR LHRH.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopitheciinae; Macaca.

OX NCBI\_Taxid=9544;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Hypothalamus;

RX MEDLINE=95124501; PubMed=7545971;

RA Ma Y.J., Costa M.E., Ojeda S.R.;

RT "Developmental expression of the genes encoding transforming growth factor alpha and its receptor in the hypothalamus of female rhesus macaques.";

RT Neuroendocrinology 60:346-359(1994).

RL -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: Belongs to the GnRH family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

RN [1]



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:00:31; Search time 6.68144 Seconds  
(without alignments)  
253.383 Million cell updates/sec

Title: US-09-848-834A-16  
Perfect score: 183  
Sequence: 1 XHWSYGLRPGSSGSLDEKIKAKMEKASSVFNVNS 36

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	97	53.0	388	1 CSP_PLARE	P26694 plasmodium
2	91	49.7	397	1 CSP_PLAFO	P19597 plasmodium
3	91	49.7	412	1 CSP_PLAFA	P02893 plasmodium
4	91	49.7	424	1 CSP_PLAFT	P13814 plasmodium
5	90	49.2	442	1 CSP_PLAFW	P08307 plasmodium
6	61.5	33.6	61	1 GONL_SHEEP	Q28588 ovis aries
7	59	32.2	63	1 GONL_MESAU	O09163 mesocricetu
8	58	31.7	67	1 GONL_MACMU	P55247 macaca mula
9	58	31.7	89	1 GONL_XENLA	P45656 xenopus lae
10	58	31.7	90	1 GONL_MOUSE	P13562 mus musculus
11	58	31.7	91	1 GONL_PIG	P49921 sus scrofa
12	58	31.7	92	1 GONL_HUMAN	P01148 homo sapien
13	58	31.7	92	1 GONL_RAT	P07490 rattus norv
14	58	31.7	92	1 GONL_TUPGB	Q95335 tupiaia glis
15	58	31.7	339	1 CSP_PLABE	P06915 plasmodium
16	58	31.7	347	1 CSP_PLABA	P23093 plasmodium
17	56.5	30.9	89	1 GON3_DICLA	O91a09 dicentrarch
18	55.5	30.3	89	1 GON3_PORNO	P51922 porichthys
19	55	30.1	94	1 GONL_HAPBU	P51918 haplochromi
20	55	30.1	367	1 CSP_PLAYO	P06914 plasmodium
21	55	30.1	721	1 THIC_SHEON	Q8eed7 shewanella
22	54	29.5	10	1 GONL_ALLMI	P37041 alligator m
23	54	29.5	92	1 GONL_CHICK	P37042 gallus gall
24	53.5	28.2	90	1 GON3_HAPBU	P45652 haplochromi
25	52.5	28.7	487	1 MMSA_BACSU	P42412 bacillus su
26	52	28.4	95	1 GONL_MORSA	O73812 morone saxa
27	52	28.4	95	1 GON1_PAGMA	P70074 pagrus majo
28	52	28.4	95	1 GON1_SPAAU	P51919 sparus aura
29	52	28.4	99	1 GONL_DICLA	Q91a10 dicentrarch
30	51.5	28.1	80	1 TIG_MYCGE	P33439 ciarias gar
31	51.5	28.1	444	1 TIG_MYCGE	P47480 mycoplasma
32	51	27.9	393	1 CSP_PLABR	P14593 plasmodium
33	51	27.9	429	1 CSP_PLAMA	P13815 plasmodium

34	51	27.9	959	1 MSH1_YEAST	P25846 saccharomyc
35	50.5	27.6	90	1 GON3_PAGMA	P51921 pagrus majo
36	50.5	27.6	90	1 GON3_SPAAU	P51923 sparus aura
37	50.5	27.6	222	1 GTA2_CHICK	Q08393 gallus gall
38	50	27.3	90	1 GON8_RANDY	Q91au2 rana dybows
39	50	27.3	223	1 GTAL_RABIT	Q08863 oryctolagus
40	50	27.3	433	1 THIC1_SULSO	Q97vj8 sulfolobus
41	50	27.3	590	1 THIC1_BACSU	P45740 bacillus su
42	50	27.3	595	1 THIC1_BACHD	Q9kb14 bacillus ha
43	50	27.3	1534	1 RREL_CANFA	Q28298 canis famli
44	49.5	27.0	482	1 GABD_ECOLI	P25526 escherichia
45	49	26.8	92	1 GON1_CAVPO	O54713 cavia porce

ALIGNMENTS

RESULT 1  
CSP\_PLARE STANDARD; PRT; 388 AA.  
AC P26694;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Circumsporozoite protein precursor (CS).  
OS Plasmodium reichenowi.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_Taxid=5854;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91201303; PubMed=2016283;  
RA Lal A.A., Goldman I.F.;  
RT "Circumsporozoite protein gene from Plasmodium reichenowi, a  
RT chimpanzee malaria parasite evolutionarily related to the human  
RT malaria parasite plasmodium falciparum.";  
RL J. Biol. Chem. 266:6686-6689(1991).  
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT  
CC SURFACE ANTIGEN ON THE SPOOROZITE (THE INFECTIVE STAGE OF THE  
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE  
CC VERTEBRATE HOST).  
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR  
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES  
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.  
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; M60972; AAA29561.1; -  
CC PIR; A39756; A39756.  
CC InterPro; IPR003067; Crcmsprzoite.  
CC InterPro; IPR000884; TSP1.  
CC Pfam; PF00090; tsp-1; 1.  
CC PRINTS; PR01303; Crcmsprzoite.  
CC SMART; SM00209; TSP1; 1.  
CC PROSITE; PS50092; TSP1; 1.  
CC Malaria; Sporozoite; Repeat; Signal.  
CC SIGNAL 1 16 PROBABLE.  
CC FT CHAIN 17 388 CIRCUMSPOROZOITE PROTEIN.  
CC FT DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.  
CC FT DOMAIN 313 366 TSP TYPE-1.  
CC SQ SEQUENCE 388 AA; 42245 MW; C031EEFBE2E35604 CRC64;  
Query Match 53.0%; Score 97; DB 1; Length 388;  
Best Local Similarity 57.9%; Pred. No. 2.7e-06;  
Matches 22; Conservative 3; Mismatches 5; Indels 8; Gaps 1;  
QY 7 LRPSSGSLD-----EKKIAKMEKASSVFNVNS 36



```

; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-770-014-55

```

```

Query Match      44.4%; Score 110; DB 4; Length 31;
Best Local Similarity 78.6%; Pred. No. 4.8e-08;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      19 LSEIRGVIVHRLGVGEGPSLHWSYGLRP 46
          |||||
Db       3 LSEIRGVIVHKLGVLFGGHWSYGLRP 30
          |||||

```

Search completed: September 24, 2003, 17:00:25  
 Job time : 15.0307 secs

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-770-014-43

Query Match 45.6%; Score 113; DB 4; Length 27;
Best Local Similarity 82.1%; Pred. No. 1.6e-08;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 19 LSEIKGVIVHRLEGVPSLHWSYGLRP 46
   |||||:|||||
Db 1 LSEIKGVIVHKLEGVGGE--HWSYGLRP 26

RESULT 13
US-09-100-414B-55
; Sequence 55, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-55

Query Match 44.4%; Score 110; DB 3; Length 31;
Best Local Similarity 78.6%; Pred. No. 4.8e-08;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGVPSLHWSYGLRP 46
   |||||:|||||
Db 3 LSEIKGVIVHKLEGVGGEHWSYGLRP 30

RESULT 15
US-09-770-014-55
; Sequence 55, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; CORRESPONDENCE ADDRESS:
```

```

; ADDRESS: Morgan & Finnegan, L.L.P.
;
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:

```

```
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; *SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-770-014-36

Query Match 46.8%; Score 116; DB 4; Length 27;
Best Local Similarity 85.7%; Pred. No. 6.4e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
Db 1 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 26

RESULT 8
US-08-446-692-33
; Sequence 33, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-33

Query Match 46.8%; Score 116; DB 2; Length 45;
Best Local Similarity 85.7%; Pred. No. 1.2e-08;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
```

```
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-33

Query Match 46.8%; Score 116; DB 1; Length 45;
Best Local Similarity 85.7%; Pred. No. 1.2e-08;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
Db 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 44

RESULT 9
US-08-488-351A-33
; Sequence 33, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-33

Query Match 46.8%; Score 116; DB 2; Length 45;
Best Local Similarity 85.7%; Pred. No. 1.2e-08;
```

; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/229,275  
 ; FILING DATE: 14-APR-1994  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/057,166  
 ; FILING DATE: 27-APR-1992  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Maria C.H. Lin  
 ; REGISTRATION NUMBER: 29,323  
 ; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)415-8745  
 ; TELEFAX: (516)751-6849  
 ; INFORMATION FOR SEQ ID NO: 19:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 27 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-488-351A-19

Query Match 46.8%; Score 116; DB 2; Length 27;  
 Best Local Similarity 85.7%; Pred. No. 6.4e-09;  
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 46  
 Db 1 LSEIKGVIVHRLEGVGPGLHWSYGLRP 26

RESULT 5  
 US-09-100-414B-36  
 ; Sequence 36, Application US/09100414B  
 ; Patent No. 6025468  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Chang Yi  
 ; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
 ; TITLE OF INVENTION: IMMUNOGENS  
 ; NUMBER OF SEQUENCES: 106  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Morgan & Finnegan, L.L.P.  
 ; STREET: 345 Park Avenue  
 ; CITY: New York  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10154-0054  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC Windows  
 ; SOFTWARE: Word 97  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/100.414B  
 ; FILING DATE: 20-JUNE-1998  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Maria H. Lin  
 ; REGISTRATION NUMBER: 29,323  
 ; REFERENCE/DOCKET NUMBER: 1151-4157  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-758-4800  
 ; TELEFAX: 212-751-6849  
 ; INFORMATION FOR SEQ ID NO: 36:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 27 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-09-100-414B-36

Query Match 46.8%; Score 116; DB 3; Length 27;

Best Local Similarity 85.7%; Pred. No. 6.4e-09;  
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;  
 QY 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 46  
 Db 1 LSEIKGVIVHRLEGVGPGLHWSYGLRP 26

RESULT 6  
 US-09-303-323-36  
 ; Sequence 36, Application US/09303323  
 ; Patent No. 6228987  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Chang Yi  
 ; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
 ; TITLE OF INVENTION: IMMUNOGENS  
 ; NUMBER OF SEQUENCES: 106  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Morgan & Finnegan, L.L.P.  
 ; STREET: 345 Park Avenue  
 ; CITY: New York  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10154-0054  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC Windows  
 ; SOFTWARE: Word 97  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/303.323  
 ; FILING DATE: 30-APR-1999  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/100.414  
 ; FILING DATE: 20-JUNE-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Maria H. Lin  
 ; REGISTRATION NUMBER: 29,323  
 ; REFERENCE/DOCKET NUMBER: 1151-4157  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-758-4800  
 ; TELEFAX: 212-751-6849  
 ; INFORMATION FOR SEQ ID NO: 36:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 27 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-09-303-323-36

Query Match 46.8%; Score 116; DB 3; Length 27;  
 Best Local Similarity 85.7%; Pred. No. 6.4e-09;  
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 46  
 Db 1 LSEIKGVIVHRLEGVGPGLHWSYGLRP 26

RESULT 7  
 US-09-770-014-36  
 ; Sequence 36, Application US/09770014  
 ; Patent No. 6559282  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Chang Yi  
 ; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
 ; TITLE OF INVENTION: IMMUNOGENS  
 ; NUMBER OF SEQUENCES: 106  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Morgan & Finnegan, L.L.P.  
 ; STREET: 345 Park Avenue  
 ; CITY: New York

```

; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-19

Query Match 46.8%; Score 116; DB 1; Length 27;
Best Local Similarity 85.7%; Pred. No. 6.4e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGP SLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | |
Db 1 LSEIKGVIVHRLEGVGP --HWSYGLRP 26

RESULT 4
US-08-488-351A-19
; Sequence 19, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488.351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446.692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: US 08/229.275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057.166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-20

Query Match 47.2%; Score 117; DB 2; Length 42;
Best Local Similarity 82.8%; Pred. No. 8.1e-09;
Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 18 LLEIKGVIVHRLEGVGP SLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | |
Db 15 VLSEIKGVIVHRLEGVGP --HWSYGLRP 41

RESULT 3
US-08-446-692-19
; Sequence 19, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:46:20 ; Search time 13.9307 Seconds  
(without alignments)  
142.750 Million cell updates/sec

Title: US-09-848-834A-17  
Perfect score: 248  
Sequence: 1 XHWSYGLRPGSGPSLKLLS.....HRLGEGVPSLHWSYGLRPX 47

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	47.2	42	1 US-08-446-692-20	Sequence 20, Appl
2	117	47.2	42	2 US-08-488-351A-20	Sequence 20, Appl
3	116	46.8	27	1 US-08-446-692-19	Sequence 19, Appl
4	116	46.8	27	2 US-08-488-351A-19	Sequence 19, Appl
5	116	46.8	27	3 US-09-100-414B-36	Sequence 36, Appl
6	116	46.8	27	4 US-09-303-323-36	Sequence 36, Appl
7	116	46.8	27	4 US-09-770-014-36	Sequence 36, Appl
8	116	46.8	45	1 US-08-446-692-33	Sequence 33, Appl
9	116	46.8	45	2 US-08-488-351A-33	Sequence 33, Appl
10	113	45.6	27	3 US-09-100-414B-43	Sequence 43, Appl
11	113	45.6	27	3 US-09-303-323-43	Sequence 43, Appl
12	113	45.6	27	4 US-09-770-014-43	Sequence 43, Appl
13	110	44.4	31	3 US-09-100-414B-55	Sequence 55, Appl
14	110	44.4	31	3 US-09-303-323-55	Sequence 55, Appl
15	110	44.4	31	3 US-09-770-014-55	Sequence 55, Appl
16	108	43.5	27	3 US-09-100-414B-41	Sequence 41, Appl
17	108	43.5	27	3 US-09-303-323-41	Sequence 41, Appl
18	108	43.5	27	3 US-09-770-014-41	Sequence 41, Appl
19	108	43.5	27	4 US-09-303-323-47	Sequence 47, Appl
20	108	43.5	27	4 US-09-770-014-41	Sequence 41, Appl
21	108	43.5	27	4 US-09-770-014-47	Sequence 47, Appl
22	108	43.5	45	3 US-09-100-414B-45	Sequence 45, Appl
23	108	43.5	45	3 US-09-303-323-45	Sequence 45, Appl
24	108	43.5	45	4 US-09-770-014-45	Sequence 45, Appl
25	107	43.1	31	3 US-09-100-414B-59	Sequence 59, Appl
26	107	43.1	31	3 US-09-303-323-59	Sequence 59, Appl
27	107	43.1	31	4 US-09-770-014-59	Sequence 59, Appl

28	106	42.7	28	3 US-09-100-414B-38	Sequence 38, Appl
29	106	42.7	28	3 US-09-303-323-38	Sequence 38, Appl
30	106	42.7	28	4 US-09-770-014-38	Sequence 38, Appl
31	106	42.7	40	2 US-08-460-502-10	Sequence 10, Appl
32	106	42.7	40	4 US-09-148-711A-10	Sequence 10, Appl
33	105	42.3	31	3 US-09-100-414B-53	Sequence 53, Appl
34	105	42.3	31	3 US-09-303-323-53	Sequence 53, Appl
35	105	42.3	31	4 US-09-770-014-53	Sequence 53, Appl
36	105	42.3	47	3 US-09-100-414B-60	Sequence 60, Appl
37	105	42.3	47	3 US-09-303-323-60	Sequence 60, Appl
38	105	42.3	47	4 US-09-770-014-60	Sequence 60, Appl
39	105	42.3	49	3 US-09-100-414B-57	Sequence 57, Appl
40	105	42.3	49	3 US-09-303-323-57	Sequence 57, Appl
41	105	42.3	49	4 US-09-770-014-57	Sequence 57, Appl
42	103.5	41.7	25	1 US-08-446-692-17	Sequence 17, Appl
43	103.5	41.7	25	2 US-08-488-351A-17	Sequence 17, Appl
44	103	41.5	27	3 US-09-100-414B-50	Sequence 50, Appl
45	103	41.5	27	3 US-09-303-323-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1  
US-08-446-692-20  
; Sequence 20, Application US/08446692  
; Patent No. 5759551  
; GENERAL INFORMATION:  
; APPLICANT: Ladd, Anna  
; APPLICANT: Wang, Chang Yi  
; APPLICANT: Zamb, Timothy  
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
; TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maria C.H. Lin  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: US  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,692  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria C.H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)415-8745  
; TELEFAX: (516)751-6849  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-446-692-20

Query Match 47.2%; Score 117; DB 1; Length 42;  
Best Local Similarity 82.8%; Pred. No. 8.1e-09;  
Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

OY 18 LLSEIKGVIVHRLGEGVPSLHWSYGLRP 46  
DB 15 VLSEIKGVIVHRLGEGVPSLHWSYGLRP 41

KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;  
KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;  
KW cholesteryl ester transport protein; anti-arteriosclerotic.  
XX  
OS Chimeric - Measles virus.  
OS Chimeric - Rattus sp.  
XX  
XX WO9966957-A2.  
XX  
XX 29-DEC-1999.  
PD  
XX  
XX 21-JUN-1999; 99WO-US13975.  
PF  
XX  
XX 20-JUN-1998; 98US-0100412.  
PR  
XX  
XX (UNBI-) UNITED BIOMEDICAL INC.  
PA  
XX  
XX Wang CY;  
PI  
XX  
XX WPI; 2000-160564/14.  
DR  
XX  
XX New artificial T helper cell epitope and derived immunogens with target  
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis  
PT or human immune deficiency virus -  
PT  
XX  
XX Example 1; Page 79; 129pp; English.  
PS  
XX  
XX The invention relates to novel promiscuous T helper cell epitopes (Th),  
XX and immunogenic peptides comprising the Th epitopes of the invention  
CC along with B cell epitopes. The Th epitopes and peptide immunogens  
CC containing them, are used to induce a T helper cell response,  
CC specifically against Plasmodium falciparum, cholesteryl ester transport  
CC protein (CERP) or HIV epitopes, but more generally against any pathogen,  
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and  
CC peptide immunogens may be used for prevention and/or treatment of  
CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
CC immunotherapy; for inhibition of the action of luteinising hormone  
CC releasing hormone (LHRH) for contraception, treatment of hormone-  
CC dependent cancer, prevention of boar taint in meat, and  
CC immunocastration); for promoting the growth of animals; or for  
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous  
CC Th (functional in genetically diverse subjects) into an immunogen  
CC improves capacity to induce a strong T helper cell-mediated immune  
CC response, resulting in production of antibodies against a target  
CC antigen. Th can replace carrier proteins and pathogen-derived T helper  
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope  
CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,  
CC AAY91126 and AAY91245-Y91246 represent synthetic Th epitopes based on the  
CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope  
CC from hepatitis B virus (HBV) surface antigen, and sequences  
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.  
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides  
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197  
CC is the LHRH target antigenic peptide used in these LHRH antigenic  
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic  
CC peptides comprising somatostatin and a Th epitope. Somatostatin  
CC immunogens may be used to promote growth in livestock. AAY91208 is a  
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th  
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV  
CC infection of T cells. AAY90212 is a modified version of a human IgE  
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3  
CC antigenic peptides which may be used in the treatment of allergies.  
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)  
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th  
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target  
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th  
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent  
CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a  
CC CERP peptide and a Th epitope which may be used to prevent or treat  
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257  
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and  
CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell

CC epitope which may be used as a component in an anti-HIV-1 vaccine.  
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive  
CC protein epitope from Yersinia species, and hinge spacer peptide, both of  
CC which may optionally be used in the antigenic peptides of the  
CC invention.  
XX  
XX SQ Sequence 27 AA;  
Query Match 43.5%; Score 108; DB 21; Length 27;  
Best Local Similarity 71.4%; Pred. No. 3e-07; 2; Gaps 1;  
Matches 20; Conservative 4; Mismatches 2; Indels 2;  
QY 19 LSEIKGVIVHRLEGVEGPGSLHWSYGLRP 46  
:|||||||:|:|:| | |||||  
Db 1 ISEIKGVIVHKIEGIGGE--HWSYGLRP 26  
Search completed: September 24, 2003, 17:40:08  
Job time : 54.1191 secs

infection of T cells. AAY90212 is a modified version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3 antigenic peptides which may be used in the treatment of allergies. AAY91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th epitope, and may be used in a malaria vaccine. AAY91228-Y91231 represent CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a CERP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVF Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory invasive protein epitope from *Yersinia* species, and hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.

Query Match 45.6%; Score 113; DB 21; Length 27;  
 Best Local Similarity 82.1%; Pred. No. 6.6e-08;  
 Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;  
 QY 19 LSEIKGVIVHRLGEGVPSLHWSYGLRP 46  
 DB 1 LSEIKGVIVHRLGEGVGE--HWSYGLRP 26

RESULT 14  
 AAY91175  
 ID AAY91175 standard; peptide; 31 AA.

AC AAY91175;

DT 22-MAY-2000 (first entry)

DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:55.

KW Promiscuous T-cell epitope; measles virus F protein; MVF;  
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;  
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;  
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;  
 KW cholesterol ester transport protein; anti-arteriosclerotic.

OS Chimeric - Measles virus.  
 OS Chimeric - Rattus sp.

PN W09966957-A2.

PD 29-DEC-1999.

PF 21-JUN-1999; 99WO-US13975.

PR 20-JUN-1998; 98US-0100412.

PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY;

DR WPI; 2000-160564/14.

PT New artificial T helper cell epitope and derived immunogens with target antigenic site, for immunization against e.g. malaria, arteriosclerosis or human immune deficiency virus -

PS Example 1; Page 84; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),  
 CC and immunogenic peptides comprising the Th epitopes of the invention

CC along with B cell epitopes. The Th epitopes and peptide immunogens  
 CC containing them, are used to induce a T helper cell response,  
 CC specifically against Plasmodium falciparum, cholesterol ester transport  
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,  
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and  
 CC peptide immunogens may be used for prevention and/or treatment of  
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
 CC immunotherapy; for inhibition of the action of luteinising hormone  
 CC releasing hormone (LHRH) for contraception, treatment of hormone-  
 CC dependent cancer, prevention of boar taint in meat, and  
 CC immunocastration; for promoting the growth of animals; or for  
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous  
 CC Th (functional in genetically diverse subjects) into an immunogen  
 CC improves capacity to induce a strong T helper cell-mediated immune  
 CC response, resulting in production of antibodies against a target  
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper  
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope  
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,  
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the  
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope  
 CC from hepatitis B virus (HBV) surface antigen, and sequences  
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.  
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides  
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197  
 CC is the LHRH target antigenic peptide used in these LHRH antigenic  
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic  
 CC peptides comprising somatostatin and a Th epitope. Somatostatin  
 CC immunogens may be used to promote growth in livestock. AAY91208 is a  
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVF Th  
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV  
 CC infection of T cells. AAY90212 is a modified version of a human IgE  
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3  
 CC antigenic peptides which may be used in the treatment of allergies.  
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)  
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th  
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target  
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th  
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent  
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a  
 CC CERP peptide and a Th epitope which may be used to prevent or treat  
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257  
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and  
 CC AAY91258-Y91273 are antigenic peptides comprising MVF Th and HIV-1 B-cell  
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.  
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive  
 CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of  
 CC which may optionally be used in the antigenic peptides of the  
 CC invention.

XX Sequence 31 AA;

Query Match 44.4%; Score 110; DB 21; Length 31;  
 Best Local Similarity 78.6%; Pred. No. 2e-07;  
 Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGEGVPSLHWSYGLRP 46

DB 3 LSEIKGVIVHRLGEGVPSLHWSYGLRP 30

RESULT 15  
 AAY91161

ID AAY91161 standard; peptide; 27 AA.

AC AAY91161;

XX 22-MAY-2000 (first entry)

DT Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:41.

DE Promiscuous T-cell epitope; measles virus F protein; MVF;

KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;



XX The invention relates to novel promiscuous T helper cell epitopes (Th),  
 XX and immunogenic peptides comprising the Th epitopes of the invention  
 CC along with B cell epitopes. The Th epitopes and peptide immunogens  
 CC containing them, are used to induce a T helper cell response,  
 CC specifically against Plasmodium falciparum, cholesterol ester transport  
 CC protein (CTP) or HIV epitopes, but more generally against any pathogen,  
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and  
 CC peptide immunogens may be used for prevention and/or treatment of  
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
 CC immunotherapy; for inhibition of the action of luteinising hormone  
 CC releasing hormone (LHRH) for contraception, treatment of hormone-  
 CC dependent cancer, prevention of boar taint in meat, and  
 CC immunocastration); for promoting the growth of animals; or for  
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous  
 CC Th (functional in genetically diverse subjects) into an immunogen  
 CC improves capacity to induce a strong T helper cell-mediated immune  
 CC response, resulting in production of antibodies against a target  
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper  
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope  
 CC from the measles virus F (WVF) protein and sequences AAY91122-Y91142,  
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the  
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope  
 CC from hepatitis B virus (HBV) surface antigen, and sequences  
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.  
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides  
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197  
 CC is the LHRH target antigenic peptide used in these LHRH antigenic  
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic  
 CC peptides comprising somatostatin and a Th epitope. Somatostatin  
 CC immunogens may be used to promote growth in livestock. AAY91208 is a  
 CC human CD4 CD82-like domain antigenic site, and AAY91209-Y90211 are MVH Th  
 CC epitope/CD4 CD82 antigenic peptides which may be used to prevent HIV  
 CC infection of T cells. AAY90212 is a modified version of a human IgE  
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3  
 CC antigenic peptides which may be used in the treatment of allergies.  
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)  
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th  
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target  
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent  
 CC CPT-derived peptides and AAY91232-Y91241 are immunogens comprising a  
 CC CPT peptide and a Th epitope which may be used to prevent or treat  
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257  
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and  
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell  
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.  
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasion  
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of  
 CC which may optionally be used in the antigenic peptides of the  
 CC invention.

XX Sequence 27 AA;

Query Match 46.8%; Score 116; DB 21; Length 27;

Best Local Similarity 85.7%; Pred. No. 2.7e-08;

Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGSLSHWSYGLRP 46  
 |||||  
 Db 1 LSEIKGVIVHRLEGVGSLSHWSYGLRP 26

RESULT 11

AAY68567

ID AAY68567 standard; peptide; 27 AA.

XX AAY68567;

XX 05-MAY-2000 (first entry)

XX

DE Peptide immunogen comprising a Th epitope and LHRH target antigen.  
 XX  
 KW Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;  
 CC luteinising hormone-releasing hormone; spermatogenesis; ovulation;  
 CC oestrus; sexual development; sex hormone; promiscuous T helper epitope;  
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;  
 KW breast cancer; endometriosis; boar taint; meat quality; chimera;  
 KW immunocastration.  
 XX  
 XX Chimeric - Measles virus.  
 OS Chimeric - Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..15  
 FT /note= "helper Th epitope AAY68540"  
 FT 16..17  
 FT /note= "spacer"  
 FT 18..27  
 FT /note= "LHRH antigenic epitope AAY68566"  
 XX  
 PN W09966952-A1.  
 XX  
 XX 29-DEC-1999.  
 PD  
 XX 21-JUN-1999; 99WO-US13960.  
 PF  
 XX 20-JUN-1998; 98US-0100414.  
 PR  
 XX (UNBI-) UNITED BIOMEDICAL INC.  
 PA  
 XX Wang CY;  
 PI  
 XX WPI; 2000-160562/14.  
 DR  
 XX New peptide immunogen containing luteinising hormone-releasing hormone  
 PT antigen site and helper T cell epitope, for e.g. contraception and  
 PT treatment of cancer -  
 PS  
 XX Example 1; Page 63; 102pp; English.  
 XX  
 CC The present sequence represents a peptide immunogen comprising a  
 CC helper T cell (Th) epitope of the F protein of the Measles virus and  
 CC a target antigen, luteinising hormone-releasing hormone (LHRH).  
 CC The peptide immunogens cause induction of a specific immune response  
 CC to LHRH which is involved in regulation of spermatogenesis, ovulation,  
 CC oestrus, sexual development and secretion of sex hormones. Provision of  
 CC a promiscuous T helper epitope (which is functional in genetically  
 CC diverse subjects) provides optimum immunogenicity to the B cell  
 CC epitopes of the target antigen and thus high antibody titres against  
 CC the target antigen. The peptide immunogens of the invention are used  
 CC to vaccinate against mammalian LHRH, for use as (reversible)  
 CC contraceptive; control of hormone-dependent tumours (cancer of prostate  
 CC or breast, also endometriosis); to prevent boar taint (and improve meat  
 CC quality) and for immunocastration.  
 XX  
 SQ Sequence 27 AA;

Query Match 46.8%; Score 116; DB 21; Length 27;

Best Local Similarity 85.7%; Pred. No. 2.7e-08;

Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGSLSHWSYGLRP 46  
 |||||  
 Db 1 LSEIKGVIVHRLEGVGSLSHWSYGLRP 26

RESULT 12

AAR62721

ID AAR62721 standard; peptide; 45 AA.

XX AAR62721;

XX 25-MAR-2003 (updated)

DT



CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-CnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.

XX Sequence 50 AA;

Query Match 57.1%; Score 141.5; DB 23; Length 50;  
 Best Local Similarity 58.8%; Pred. No. 2.4e-11;  
 Matches 30; Conservative 1; Mismatches 11; Indels 9; Gaps 2;

QY 2 HWSYGLRPGSGPSL-----KLSEIKGVIVHRLEGVGSPSLHWSYGLRP 46  
 ||||| ||||| ||||| : | ||||| |||||  
 Db 2 HWSYGLRPGSGPSLFFNFTVFWLRVPKVSASHL---EGPSLHWSYGLRP 49

#### RESULT 7

AAR62705  
 ID AAR62705 standard; peptide; 25 AA.

XX AC AAR62705;

XX 25-MAR-2003 (updated)  
 DT 10-SEP-1995 (first entry)

XX LHRH-containing immunogenic peptide.

XX Helper T cell epitope; universal immune stimulator; invasive; haptens;  
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;  
 KW androgen-dependent carcinoma; antitumour; infertility;  
 KW measles virus F protein.

XX Synthetic.

XX Key Location/Qualifiers  
 FH Domain 1..15  
 FT /note= "measles virus F protein helper T cell epitope"  
 FT Domain 16..25  
 FT /note= "LHRH haptens"

XX WO9425060-A1.

XX 10-NOV-1994.

XX 28-APR-1994; 94WO-US04832.

XX 27-APR-1993; 93US-0057166.  
 PR 14-APR-1994; 94US-0229275.

XX (LADD/) LADD A E.

PA (WANG/) WANG C Y.

PA (ZAMB/) ZAMB T.

XX Ladd AE, Wang CY, Zamb T;

XX WPI; 1994-357910/44.

XX Immunogenic luteinising hormone releasing hormone peptide(s) -  
 PT that suppress LHRH activity in males and females

XX Claim 8; Page 84; 213pp; English.

XX Synthetic immunogenic peptides are provided in which a universal immune  
 CC stimulator is linked to a peptide or protein haptens containing B cell  
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes  
 CC potent immune responses to the coupled peptide or protein. The  
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)  
 CC which elicits an immune response to the coupled peptide in members of  
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)  
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.  
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the  
 CC invasive and Th domains and between the immune stimulator and haptens

CC components. When the haptens is LHRH, then optionally the invasive domain  
 CC can be omitted from the immune stimulator component.  
 CC The present sequence represents an LHRH-containing, invasive-free  
 CC immunogenic peptide as above which can be used as a potent vaccine for  
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,  
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign  
 CC uterine tumours, recurrent functional ovarian cysts, (severe)  
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for  
 CC induction of infertility.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 25 AA;

Query Match 47.2%; Score 117; DB 15; Length 25;  
 Best Local Similarity 85.7%; Pred. No. 1.8e-08;  
 Matches 24; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGSPSLHWSYGLRP 46  
 ||||| ||||| ||||| |||||  
 Db 1 LSEIKGVIVHRLEGVGSPSLHWSYGLRP 24

#### RESULT 8

AAR62708

ID AAR62708 standard; peptide; 42 AA.

XX AC AAR62708;

XX 25-MAR-2003 (updated)  
 DT 10-SEP-1995 (first entry)

XX LHRH-containing immunogenic peptide.

XX Helper T cell epitope; universal immune stimulator; invasive; haptens;  
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;  
 KW androgen-dependent carcinoma; antitumour; infertility;  
 KW measles virus F protein.

XX Synthetic.

XX Key Location/Qualifiers  
 FH Domain 1..15  
 FT /note= "measles virus F protein helper T cell epitope"  
 FT Domain 16..30  
 FT /note= "measles virus F protein helper T cell epitope"  
 FT Domain 33..42  
 FT /note= "LHRH haptens"

XX WO9425060-A1.

XX 10-NOV-1994.

XX 28-APR-1994; 94WO-US04832.

XX 27-APR-1993; 93US-0057166.  
 PR 14-APR-1994; 94US-0229275.

XX (LADD/) LADD A E.

PA (WANG/) WANG C Y.

PA (ZAMB/) ZAMB T.

XX Ladd AE, Wang CY, Zamb T;

XX WPI; 1994-357910/44.

XX Immunogenic luteinising hormone releasing hormone peptide(s) -  
 PT that suppress LHRH activity in males and females

XX Claim 8; Page 86; 213pp; English.

XX Synthetic immunogenic peptides are provided in which a universal immune  
 CC stimulator is linked to a peptide or protein haptens containing B cell  
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes



```

OS Synthetic.
XX Key Location/Qualifiers
FH Peptide 1..18
FT Peptide /note= "Measles virus fusion protein F epitope"
FT Peptide 19..22
FT Peptide /note= "Spacer peptide"
FT Peptide 23..31
FT Modified-site 31
FT /note= "Gonadotrophin releasing hormone epitope"
FT /note= "Amidated glycine or glycinamide"
XX
PN WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US14363.
XX
PR 05-MAY-2000; 2000US-202328P.
XX (APHT-) APHTON CORP.
XX Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
DR Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomimic peptide
XX epitope or its analogue
XX
PS Claim 11; Page 7; 43pp; English.
XX
CC The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH) also known
CC as luteinising hormone releasing hormone, LHRH) comprising a fusion
CC peptide which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic
CC immunogen is useful inducing an immune response against GnRH in an
CC animal subject, and as such is useful as a contraceptive and in the
CC treatment of diseases such as cancer (of the breast, uterus and other
CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic
CC hypertrophy and prostate cancer. The immunogen is effective in eliciting
CC high and specific anti-GnRH antibody titres. The present sequence
CC is a synthetic immunogen of the invention.
XX
SQ Sequence 31 AA;
Query Match 64.1%; Score 159; DB 23; Length 31;
Best Local Similarity 100.0%; Pred. No. 6.4e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 KLLSEIKGVIIVHRLGEGVPSLHWSYGLRP 46
DB 1 KLLSEIKGVIIVHRLGEGVPSLHWSYGLRP 30
RESULT 4
AAU11430
ID AAU11430 standard; peptide; 46 AA.
XX
AC AAU11430;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 11.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX

```

```

OS Chimeric - Clostridium tetani.
OS Chimeric - Mammalia.
XX Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 1
FT /label= OTHER
FT Peptide /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT Peptide 1..10
FT /note= "Gonadotrophin releasing hormone epitope
FT (1..10 aa)"
FT Peptide 11..16
FT /note= "Spacer peptide"
FT Peptide 17..31
FT /note= "Tetanus toxoid (830-844 aa)"
FT Peptide 32..37
FT /note= "Spacer peptide"
FT Peptide 38..46
FT /note= "Gonadotrophin releasing hormone epitope
FT (2-10 aa)"
FT Modified-site 46
FT /note= "Amidated glycine or glycinamide"
XX
XX WO200185763-A2.
XX
XX 15-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US14363.
XX
XX 05-MAY-2000; 2000US-202328P.
XX (APHT-) APHTON CORP.
XX
XX Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomimic peptide
XX epitope or its analogue
XX
XX Claim 11; Page 12; 43pp; English.
XX
CC The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH) also known
CC as luteinising hormone releasing hormone, LHRH) comprising a fusion
CC peptide which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic
CC immunogen is useful inducing an immune response against GnRH in an
CC animal subject, and as such is useful as a contraceptive and in the
CC treatment of diseases such as cancer (of the breast, uterus and other
CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic
CC hypertrophy and prostate cancer. The immunogen is effective in eliciting
CC high and specific anti-GnRH antibody titres. The present sequence
CC is a synthetic immunogen of the invention.
XX
SQ Sequence 46 AA;
Query Match 59.5%; Score 147.5; DB 23; Length 46;
Best Local Similarity 62.5%; Pred. No. 3.5e-12;
Matches 30; Conservative 4; Mismatches 7; Indels 7; Gaps 2;
QY 2 HWSYGLRPGSSGSPSLKLL---SEIKGVIIVHRLGEGVPSLHWSYGLRP 46
DB 2 HWSYGLRPGSSGSPSLQYIKANSKFTGI----TELSSGSPSLHWSYGLRP 45
RESULT 5
AAU11431
ID AAU11431 standard; peptide; 51 AA.
XX
AC AAU11431;

```

FT Peptide /note= "Spacer peptide"  
FT 39..47  
FT /note= "Gonadotropin releasing hormone epitope  
FT (2-10 aa)"  
FT Modified-site  
FT 47  
FT /note= "Amidated glycine or glycinamide"  
XX  
FN WO200185763-A2.  
XX  
XX 15-NOV-2001.  
XX  
XX 04-MAY-2001; 2001WO-US14363.  
XX  
XX 05-MAY-2000; 2000US-202328P.  
XX  
XX (APHT-) APHTON CORP.  
XX  
XX Grimes S, Michaeli D, Stevens VC;  
XX WPI; 2002-049440/06.  
XX  
XX Novel synthetic immunogen for inducing immune response against  
XX gonadotropin releasing hormone, comprises fusion peptide having  
XX promiscuous helper T-cell peptide epitope and immunomimic peptide  
XX epitope or its analogue  
XX  
XX Claim 11; Page 11; 43pp; English.  
XX  
XX The invention relates to a synthetic immunogen for inducing specific  
XX antibodies against gonadotropin releasing hormone (GnRH) also known  
XX as luteinising hormone releasing hormone, LHRH) comprising a fusion  
XX peptide which comprises a promiscuous helper T-cell peptide epitope and  
XX immunomimic peptide epitope or its analogue. The synthetic  
XX immunogen is useful inducing an immune response against GnRH in an  
XX animal subject, and as such is useful as a contraceptive and in the  
XX treatment of diseases such as cancer (of the breast, uterus and other  
XX gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
XX hypertrophy and prostate cancer. The immunogen is effective in eliciting  
XX high and specific anti-GnRH antibody titres. The present sequence  
XX is a synthetic immunogen of the invention.  
SQ Sequence 47 AA;  
  
Query Match 99.2%; Score 246; DB 23; Length 47;  
Best Local Similarity 100.0%; Pred. No. 3.2e-25;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 2 HWSYGLRPGSGPSKLLSEIKGVIVHRLGEGVPSLHWSYGLRP 46  
Db 2 HWSYGLRPGSGPSKLLSEIKGVIVHRLGEGVPSLHWSYGLRP 46  
|||||  
RESULT 2  
AAU11424  
ID AAU11424 standard; peptide; 34 AA.  
XX  
XX AAU11424;  
XX  
XX 12-MAR-2002 (first entry)  
XX  
XX Synthetic immunogen peptide 5.  
XX  
XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
XX luteinising hormone releasing hormone; LHRH; contraceptive;  
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
XX  
XX Chimeric - Measles virus.  
OS Chimeric - Mammalia.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers

Misc-difference 1 /label= OTHER  
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
FT Peptide 1..10  
FT /note= "Gonadotropin releasing hormone epitope"  
FT Peptide 11..16  
FT /note= "Spacer peptide"  
FT Peptide 17..34  
FT /note= "Measles virus fusion protein F epitope"  
XX  
XX WO200185763-A2.  
XX  
XX 15-NOV-2001.  
XX  
XX 04-MAY-2001; 2001WO-US14363.  
XX  
XX 05-MAY-2000; 2000US-202328P.  
XX  
XX (APHT-) APHTON CORP.  
XX  
XX Grimes S, Michaeli D, Stevens VC;  
XX WPI; 2002-049440/06.  
XX  
XX Novel synthetic immunogen for inducing immune response against  
XX gonadotropin releasing hormone, comprises fusion peptide having  
XX promiscuous helper T-cell peptide epitope and immunomimic peptide  
XX epitope or its analogue  
XX  
XX Claim 11; Page 9; 43pp; English.  
XX  
XX The invention relates to a synthetic immunogen for inducing specific  
XX antibodies against gonadotropin releasing hormone (GnRH) also known  
XX as luteinising hormone releasing hormone, LHRH) comprising a fusion  
XX peptide which comprises a promiscuous helper T-cell peptide epitope and  
XX immunomimic peptide epitope or its analogue. The synthetic  
XX immunogen is useful inducing an immune response against GnRH in an  
XX animal subject, and as such is useful as a contraceptive and in the  
XX treatment of diseases such as cancer (of the breast, uterus and other  
XX gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
XX hypertrophy and prostate cancer. The immunogen is effective in eliciting  
XX high and specific anti-GnRH antibody titres. The present sequence  
XX is a synthetic immunogen of the invention.  
SQ Sequence 34 AA;  
  
Query Match 69.8%; Score 173; DB 23; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1e-15;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 2 HWSYGLRPGSGPSKLLSEIKGVIVHRLGEGVE 34  
Db 2 HWSYGLRPGSGPSKLLSEIKGVIVHRLGEGVE 34  
|||||  
RESULT 3  
AAU11420  
ID AAU11420 standard; peptide; 31 AA.  
XX  
XX AAU11420;  
XX  
XX 12-MAR-2002 (first entry)  
XX  
XX Synthetic immunogen peptide 1.  
XX  
XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
XX luteinising hormone releasing hormone; LHRH; contraceptive;  
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
XX  
XX Chimeric - Measles virus.  
OS Chimeric - Mammalia.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:58:36 ; Search time 53.1191 Seconds  
(without alignments)  
140.442 Million cell updates/sec

Title: US-09-848-834A-17  
Perfect score: 248  
Sequence: 1 XHWSYGLRPGSGPSLKLS.....HRLGEGVPSLHWSYGLRPX 47

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*  
1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	246	99.2	47	AAU11428	Synthetic immunoge
2	173	69.8	34	AAU11424	Synthetic immunoge
3	159	64.1	31	AAU11420	Synthetic immunoge
4	147.5	59.5	46	AAU11430	Synthetic immunoge
5	144	58.1	51	AAU11431	Synthetic immunoge
6	141.5	57.1	50	AAU11429	Synthetic immunoge
7	117	47.2	25	AA62705	LHRH-containing im
8	117	47.2	42	AA62708	LHRH-containing im
9	116	46.8	27	AA62707	LHRH-containing im

10	116	46.8	27	21	AA91156	MVF Th epitope/LHR
11	116	46.8	27	21	AA91156	Peptide immunogen
12	116	46.8	45	15	AA62721	LHRH-containing im
13	113	45.6	27	21	AA91163	Modified MVF Th ep
14	110	44.4	31	21	AA91175	Modified MVF Th ep
15	108	43.5	27	21	AA91161	Modified MVF Th ep
16	108	43.5	27	21	AA91167	Modified MVF Th ep
17	108	43.5	45	21	AA91165	Modified MVF Th ep
18	108	43.5	45	21	AA68573	Peptide immunogen
19	107	43.1	31	21	AA91179	Modified MVF Th ep
20	106	42.7	28	21	AA91158	Modified MVF Th ep
21	106	42.7	40	20	AA67581	Synthetic chimera f
22	106	42.7	40	21	AA91170	Measles virus fusi
23	105	42.3	31	21	AA91173	Modified MVF Th ep
24	105	42.3	31	21	AA68582	Peptide immunogen
25	105	42.3	47	21	AA91180	Inv epitope/modifi
26	105	42.3	47	21	AA68583	Peptide immunogen
27	105	42.3	49	21	AA91177	Modified MVF Th ep
28	103	41.5	27	21	AA91170	Modified MVF Th ep
29	103	41.5	27	21	AA68575	Peptide immunogen
30	103	41.5	35	21	AA91242	Modified MVF Th ep
31	102	41.1	75	24	ABP7235	Frizzled putative
32	101	40.7	40	20	AA67582	Synthetic chimera f
33	101	40.7	75	24	ABP7236	Frizzled putative
34	100	40.3	28	15	AA62726	LHRH-containing im
35	100	40.3	46	15	AA62728	LHRH-containing im
36	100	40.3	47	21	AA91183	Inv epitope/modifi
37	100	40.3	47	21	AA68586	Peptide immunogen
38	99	39.9	28	21	AA91159	Modified MVF Th ep
39	96.5	38.9	42	21	AA620865	GnRH tandem dimer
40	96.5	38.9	49	17	AAW03944	GnRH 4-repeat sequ
41	96.5	38.9	49	19	AAW03944	GnRH-2. Synthetic
42	96.5	38.9	49	19	AAW03944	Peptide hormone Gn
43	96.5	38.9	49	21	AAW61542	Four-copy gonadotr
44	96.5	38.9	49	21	AAW58135	GnRH analogue mult
45	96.5	38.9	544	17	AAW03943	LKT-GnRH protein f

ALIGNMENTS

RESULT 1  
AAU11428  
ID AAU11428 standard; peptide; 47 AA.  
XX  
AC AAU11428;  
XX  
XX 12-MAR-2002 (first entry)  
DT  
DE Synthetic immunogen peptide 9.  
XX  
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
KW Lutetinsing hormone releasing hormone; LHRH; contraceptive;  
KW promiscuous helper T-cell peptide epitope; immunomic peptide epitope;  
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
XX  
XX Chimeric - Plasmodium falciparum.  
OS Chimeric - Mammalia.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1 /label= OTHER  
FT Peptide /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
FT Peptide /note= "Gonadotrophin releasing hormone epitope (1..10 aa)"  
FT Peptide /note= "Spacer peptide"  
FT Peptide /note= "Malaria CSP protein (288-302 aa)"  
FT Peptide /note= "Malaria CSP protein (288-302 aa)"



C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>  
F:89-106/Domain: transmembrane #status predicted <TM1>  
F:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>  
F:194-219/Domain: transmembrane #status predicted <TM2>  
F:575-595/Domain: transmembrane #status predicted <TM3>  
F:110,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.2%; Score 65; DB 1; Length 631;  
Best Local Similarity 68.4%; Pred. No. 4.4;  
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLLSEIKGVIVHRLGV 33  
| | | | | | | | | | | | | | | |  
Db 365 SYPTLSEVKGVVHRLVAV 383

Search completed: September 24, 2003, 17:49:35  
Job time : 17.5346 secs

## VGNZRL

cell fusion glycoprotein precursor - rinderpest virus (strain L)  
 N:Contains: fusion glycoprotein F1; fusion glycoprotein F2  
 C:Species: rinderpest virus  
 C>Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 16-Jul-1999  
 C:Accession: A28921  
 R:Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.  
 Virology 164, 523-530, 1988  
 A:Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of the  
 A:Reference number: A28921; MUID:88219541; PMID:3285575  
 A:Accession: A28921  
 A:Molecule type: mRNA  
 A:Residues: 1-546 <TSU>  
 A:Cross-references: GB:M20870; NID:g333898; PIDN:AAA47399.1; PID:g333899  
 C:Genetics:  
 A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein

C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>  
 F:105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>  
 F:109-133/Domain: transmembrane #status predicted <TN1>  
 F:485-513/Domain: transmembrane #status predicted <TN2>  
 F:25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 66; DB 1; Length 546;  
 Best Local Similarity 93.3%; Pred. No. 2.8;  
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33

Db 284 LSEIKGVIVHRLGV 298

## RESULT 12

## VGNZCD

cell fusion glycoprotein precursor - canine distemper virus  
 N:Contains: fusion protein F1; fusion protein F2  
 C:Species: canine distemper virus  
 C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Jul-1999  
 C:Accession: J50321  
 R:Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.  
 Virus Res. 8, 373-386, 1987  
 A:Title: The nucleotide sequence of the gene encoding the F protein of canine distemper  
 A:Reference number: J50321; MUID:88129050; PMID:3433924  
 A:Accession: J50321  
 A:Molecule type: mRNA  
 A:Residues: 1-662 <BAR>  
 A:Cross-references: GB:M21849; NID:g323241; PIDN:AAA42878.1; PID:g323242  
 C:Genetics:  
 A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein  
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-135/Domain: signal sequence #status predicted <SIG>  
 F:136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>  
 F:225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>  
 F:606-629/Domain: transmembrane #status predicted <MEN>  
 F:62,141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 66; DB 1; Length 662;  
 Best Local Similarity 73.7%; Pred. No. 3.5;  
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLLSEIKGVIVHRLGV 33

Db 396 SYPTLSEVKGIVVHRLVAV 414

## RESULT 13

## S21382

cell fusion protein - canine distemper virus  
 C:Species: canine distemper virus  
 C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Nov-1999

C:Accession: S21382

R:Wild, T.P.; Bernard, A.; Spehner, D.; Villevail, D.; Drillien, R.

submitted to the EMBL Data Library, April 1992

A:Description: Vaccination of mice against canine distemper virus induced encephalitis

A:Reference number: S21382

A:Accession: S21382

A>Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-662 <WTL>

A:Cross-references: EMBL:X65509; NID:g58853; PIDN:CAA46481.1; PID:g58854

C:Superfamily: parainfluenza virus cell fusion protein

Query Match 26.6%; Score 66; DB 2; Length 662;  
 Best Local Similarity 73.7%; Pred. No. 3.5;  
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLLSEIKGVIVHRLGV 33

Db 396 SYPTLSEVKGIVVHRLVAV 414

## RESULT 14

## VGNZPD

cell fusion glycoprotein precursor - phocine distemper virus  
 N:Contains: fusion protein F1; fusion protein F2  
 C:Species: phocine distemper virus  
 C>Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 25-Oct-1996  
 C:Accession: JQ1368  
 R:Koemanes, J.; Blixenkron-Moeller, M.; Sharma, B.; Oervell, C.; Norrby, E.  
 J. Gen. Virol. 72, 2959-2966, 1991  
 A:Title: The nucleotide sequence and deduced amino acid composition of the haemagglut  
 A:Reference number: JQ1368; MUID:92113538; PMID:1765768  
 A:Accession: JQ1368  
 A:Molecule type: genomic RNA  
 A:Residues: 1-631 <KOV>

C:Genetics:

A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein

C:Keywords: glycoprotein; membrane fusion; transmembrane protein

F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>

F:89-106/Domain: transmembrane #status predicted <TM1>

F:189-193/Region: cleavage processing #status predicted

F:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>

F:194-212/Domain: transmembrane #status predicted <TM2>

F:575-595/Domain: transmembrane #status predicted <TM3>

F:110,142,148,486/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.2%; Score 65; DB 1; Length 631;  
 Best Local Similarity 68.4%; Pred. No. 4.4;  
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLLSEIKGVIVHRLGV 33

Db 365 SYPTLSEVKGIVVHRLVAV 383

## RESULT 15

## A48346

cell fusion glycoprotein precursor - phocine distemper virus (strain Ulster/88)  
 N:Contains: fusion protein F1; fusion protein F2  
 C:Species: phocine distemper virus  
 C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 25-Oct-1996  
 C:Accession: A48346  
 R:Curran, M.D.; Lu, Y.J.; Rima, B.K.

Arch. Virol. 126, 159-169, 1992

A:Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced a

A:Reference number: A48346; MUID:92398437; PMID:1524494

A:Accession: A48346

A:Molecule type: mRNA

A:Residues: 1-631 <CUR>

A>Note: sequence extracted from NCBI backbone (NCBIN:113098, NCBIP:113099)

C:Genetics:

A:Gene: F

C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-110/Product: cell fusion glycoprotein F2 #status predicted <FF2>  
F:111-553/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
F:501-517/Domain: transmembrane #status predicted <TMN>  
F:32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.0%; Score 72; DB 1; Length 553;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33  
|||||:|||||  
Db 291 LSEIKGVIVHRLEGV 305

RESULT 7  
VGNZRK  
cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)  
N:Contains: fusion glycoprotein F1; fusion glycoprotein F2  
C:Species: rinderpest virus  
C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 25-Oct-1996  
C:Accession: A31051  
R:Hsu, D.; Yamanaka, M.; Miller, J.; Dale, B.; Grubman, M.; Yilma, T.  
Virology 166, 149-153, 1988  
A:Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis  
A:Reference number: A31051; MUID:88322864; PMID:3411393  
A:Accession: A31051  
A:Molecule type: genomic RNA  
A:Residues: 1-546 <HSU>  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-108/Product: cell fusion glycoprotein F2 #status predicted <FF1>  
F:109-546/Product: cell fusion glycoprotein F1 #status predicted <FF2>  
F:109-134/Domain: transmembrane #status predicted <TMN>  
F:491-513/Domain: transmembrane #status predicted <TM2>  
F:25,57,63,518/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 28.6%; Score 71; DB 1; Length 546;  
Best Local Similarity 93.3%; Pred. No. 0.68;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33  
|||||:|||||  
Db 284 LSEIKGVIVHRLEGV 298

RESULT 8  
S47305  
gene F protein - rinderpest virus  
C:Species: rinderpest virus  
C:Date: 20-Oct-1994 #sequence\_revision 08-Sep-1995 #text\_change 20-Sep-1999  
C:Accession: S47305; S47301  
R:Baron, M.D.; Barrett, T.  
submitted to the EMBL Data Library, March 1994  
A:Description: The sequence of the N and L genes of Rinderpest virus, and the 50 and 30  
A:Reference number: S47283  
A:Accession: S47305  
A:Molecule type: mRNA  
A:Residues: 1-546 <BAR>  
A:Cross-references: EMBL:230697; NID:g535396; PIDN:CAA83181.1; PID:g535401; EMBL:230700;  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: transmembrane protein

Query Match 28.6%; Score 71; DB 2; Length 546;  
Best Local Similarity 93.3%; Pred. No. 0.68;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33  
|||||:|||||

Db 284 LSEIKGVIVHRLEGV 298

RESULT 9  
S47034  
cell fusion protein precursor - porpoise morbillivirus  
N:Alternate names: F protein  
C:Species: porpoise morbillivirus  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Nov-1999  
C:Accession: S47034  
R:Boit, G.; Gottschalk, E.; Blixenkrone-Woeller, M.; Wisaupt, R.G.A.; Welsh, M.;  
submitted to the EMBL Data Library, July 1994  
A:Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbilli  
A:Reference number: S47034  
A:Accession: S47034  
A:Molecule type: mRNA  
A:Residues: 1-552 <BO>  
A:Cross-references: EMBL:X80757; NID:g520639; PIDN:CAA56731.1; PID:g520640  
A:Experimental source: isolate Ulster 88  
A:Note: the source is designated as Cetacean morbillivirus  
C:Superfamily: parainfluenza virus cell fusion protein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-552/Product: fusion protein #status predicted <MAT>

Query Match 26.8%; Score 66.5; DB 2; Length 552;  
Best Local Similarity 61.5%; Pred. No. 2.5;  
Matches 16; Conservative 1; Mismatches 2; Indels 7; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGFSLHWSYGL 44  
|||||:|||||  
Db 290 LSEIKGVIVHRLEAV-----SYNL 308

RESULT 10  
JQ2223  
cell fusion protein F0 precursor - phocine distemper virus  
N:Contains: F1 and F2 chains  
C:Species: phocine distemper virus  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 24-Nov-1999  
C:Accession: JQ2223  
R:Visser, I.K.G.; van der Heijden, R.W.J.; van de Bilt, M.W.G.; Kenter, M.J.H.; Oever  
J. Gen. Virol. 74, 1989-1994, 1993  
A:Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites  
e virus entity.  
A:Reference number: JQ2223; MUID:93389459; PMID:8376973  
A:Accession: JQ2223  
A:Molecule type: mRNA  
A:Residues: 1-542 <VIS>  
A:Cross-references: GB:L07075  
A:Note: the authors translated the codon ATC for residue 4 as Leu  
C:Comment: This fusion protein F0 is cleaved into F1 and F2 chains.  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:16-542/Product: fusion protein #status predicted <MAT>  
F:16-99/Product: F2 chain #status predicted <F2C>  
F:105-542/Product: F1 chain #status predicted <F1C>  
F:105-135/Region: hydrophobic  
F:486-512/Domain: transmembrane #status predicted <TMN>  
F:21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 66; DB 2; Length 542;  
Best Local Similarity 73.7%; Pred. No. 2.8;  
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLLSEIKGVIVHRLEGV 33  
|||||:|||||  
Db 276 SYPTLSEIKGVIVHRLEAV 294

RESULT 11

RESULT 3  
JU0274  
cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain X  
N:Contains: fusion glycoprotein F1; fusion glycoprotein F2  
C:Species: subacute sclerosing panencephalitis virus, SSPEV  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jun-2000  
C:Accession: JU0274  
R:Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.  
Virus Genes 4, 173-181, 1990  
A:Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe  
A:Reference number: JU0274; MUID:90385702; PMID:1698327  
A:Accession: JU0274  
A:Molecule type: mRNA  
A:Residues: 1-534 <KOM>  
A:Cross-references: EMBL:D10548; NID:g222256; PIDN:BAA01405.1; PID:g222257  
A:Note: the authors translated the codon GTA for residue 459 as Gly and GGG for residue  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>  
F:108-534/Product: cell fusion glycoprotein F1 #status predicted <FF2>  
F:498-514/Domain: transmembrane #status predicted <TMN>  
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 29.0%; Score 72; DB 1; Length 534;  
Best Local Similarity 100.0%; Pred. No. 0.5;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 19 LSEIKGVIVHRLGV 33  
Db 288 LSEIKGVIVHRLGV 302  
RESULT 4  
S47300  
gene F protein - rinderpest virus  
C:Species: rinderpest virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Oct-1999  
C:Accession: S47300; PQ0865  
R:Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.  
submitted to the EMBL Data Library, March 1994  
A:Description: The complete nucleotide sequence of the fusion protein gene of the vaccin  
A:Reference number: S47299  
A:Accession: S47300  
A:Molecule type: DNA  
A:Residues: 1-546 <EVA>  
A:Cross-references: EMBL:Z31656; NID:g535406; PIDN:CAA83482.1; PID:g535407  
R:Chamberlain, R.W.; Wanwayi, H.M.; Hockley, E.; Shaila, M.S.; Goatley, L.; Knowles, N.J  
J. Gen. Virol. 74, 2775-2780, 1993  
A:Title: Evidence for different lineages of rinderpest virus reflecting their geographid  
A:Reference number: PQ0865; MUID:94103786; PMID:8277286  
A:Accession: PQ0865  
A:Molecule type: mRNA  
A:Residues: 86-191 <CHA>  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
Query Match 29.0%; Score 72; DB 2; Length 546;  
Best Local Similarity 100.0%; Pred. No. 0.51;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 19 LSEIKGVIVHRLGV 33  
Db 284 LSEIKGVIVHRLGV 298  
RESULT-5

E48556  
cell fusion glycoprotein precursor - measles virus (strain AIK-C)  
C:Species: measles virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 16-Jul-1999  
C:Accession: E48556  
R:Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.  
Virus Genes 7, 67-81, 1993  
A:Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the AIK  
A:Reference number: A48556; MUID:93227570; PMID:8470368  
A:Accession: E48556  
A:Molecule type: genomic RNA  
A:Residues: 1-550 <MOR>  
A:Cross-references: GB:S58435; NID:g299460; PIDN:AAB26145.1; PID:g299465  
A:Note: sequence extracted from NCBI backbone (NCBIN:129264, NCBIP:129272)  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>  
F:108-550/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
F:113-138/Region: hydrophobic  
F:495-514/Domain: transmembrane #status predicted <TMN>  
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 29.0%; Score 72; DB 1; Length 550;  
Best Local Similarity 100.0%; Pred. No. 0.51;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 19 LSEIKGVIVHRLGV 33  
Db 288 LSEIKGVIVHRLGV 302  
RESULT 6  
VGNZMV  
cell fusion glycoprotein precursor - measles virus  
C:Species: measles virus  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1989 #text\_change 16-Jun-2000  
C:Accession: A26962; A25616; PQ0380; PQ0384  
R:Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.  
J. Gen. Virol. 68, 1695-1703, 1987  
A:Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and co  
A:Reference number: A92794; MUID:87224816; PMID:3585281  
A:Accession: A26962  
A:Molecule type: mRNA  
A:Residues: 1-553 <BUC>  
A:Cross-references: GB:D00090; NID:g222061; PIDN:BAA00056.1; PID:g222062  
A:Experimental source: strain Halle  
R:Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Englund, G.; Bellini  
Virology 155, 508-523, 1986  
A:Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles v  
A:Reference number: A94350; MUID:87071668; PMID:3788062  
A:Accession: A25616  
A:Molecule type: mRNA  
A:Residues: 4-553 <RIC>  
A:Cross-references: GB:M14915; NID:g331762; PIDN:AAA46423.1; PID:g331763  
A:Experimental source: strain Edmonston  
R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.  
J. Gen. Virol. 73, 1581-1586, 1992  
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compar  
A:Reference number: PQ0374; MUID:92300360; PMID:1607874  
A:Accession: PQ0380  
A:Molecule type: genomic RNA  
A:Residues: 272-553 <SCH1>  
A:Experimental source: isolate CL  
A:Accession: PQ0384  
A:Molecule type: genomic RNA  
A:Residues: 272-553 <SCH2>  
A:Experimental source: isolate SE  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:30:06 ; Search time 16.5346 Seconds  
(without alignments)  
273.361 Million cell updates/sec

Title: US-09-848-834A-17

Perfect score: 248

Sequence: 1 XHWSYGLRPGSSGPSKLLS.....HRLEGVQPSLHWSYGLRPX 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	29.0	282	2 P00376	cell fusion glycop
2	72	29.0	282	2 P00388	cell fusion glycop
3	72	23.0	534	1 J00274	cell fusion glycop
4	72	29.0	546	2 S47300	gene F protein - r
5	72	29.0	550	1 E48556	cell fusion glycop
6	72	29.0	553	1 VGNZMV	cell fusion glycop
7	71	28.6	546	1 VGNZRK	cell fusion glycop
8	71	28.6	546	2 S47305	gene F protein - r
9	66.5	26.8	552	2 S47034	cell fusion protei
10	66	26.6	542	2 J02223	cell fusion protei
11	66	26.6	546	1 VGNZRL	cell fusion glycop
12	66	26.6	662	1 VGNZCD	cell fusion glycop
13	66	26.6	662	2 S21382	cell fusion protei
14	65	26.2	631	1 VGNZPD	cell fusion glycop
15	65	26.2	631	1 A48346	cell fusion glycop
16	64	25.8	67	2 I78541	gonadoliberin prec
17	64	25.8	92	1 RHUG	gonadoliberin prec
18	63.5	25.6	89	2 I50739	gonadoliberin prec
19	62.5	25.2	88	2 I51423	gonadoliberin prec
20	60.5	24.4	82	2 I51365	gonadoliberin-relea
21	60.5	24.4	555	2 B72486	gonadoliberin-relea
22	60.5	24.4	880	2 B53743	probable hydantoin
23	60	24.2	546	2 S5386	protein-tyrosine k
24	59.5	24.0	90	1 RMSG	cell fusion protei
25	59.5	24.0	1041	2 C83548	gonadoliberin prec
26	58.5	23.6	502	2 T36589	hypothetical prote
27	58	23.4	10	1 RHPGG	probable transmem
28	58	23.4	10	1 RHSHG	gonadoliberin - sh
29	58	23.4	92	1 RHRTG	gonadoliberin prec

30	57.5	23.2	509	2 T45529	agaA protein [impo
31	57.5	23.2	880	1 JC4166	protein-tyrosine k
32	57	23.0	371	2 T49908	hypothetical prote
33	56.5	22.8	74	2 I51092	gonadotropin relea
34	56.5	22.8	82	2 I51355	gonadotropin relea
35	56.5	22.8	82	2 I51331	gonadotropin relea
36	56	22.6	233	2 E87362	hypothetical prote
37	56	22.6	233	2 S47299	gene F protein - r
38	55.5	22.4	233	2 H69021	tetrahydromethanop
39	55.5	22.4	233	2 S38366	gonadoliberin I pr
40	55	22.2	92	2 I50644	phosphoribosylami
41	55	22.2	379	1 DCBSPK	probable galactosi
42	55	22.2	400	2 AI0104	aldehyde reductase
43	54.5	22.0	80	2 S39779	probable tetr-fami
44	54.5	22.0	190	2 T37168	36K antigen pra -
45	54.5	22.0	249	2 A41497	

ALIGNMENTS

RESULT 1

P00376

cell fusion glycoprotein - measles virus (strain TT) (fragment)

C:Species: measles virus

C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 24-Nov-1999

C:Accession: P00376

R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.

J. Gen. Virol. 73, 1581-1586, 1992

A>Title: A measles virus isolate from a child with Kawasaki disease: sequence compari

A:Reference number: P00374; MUID:92300360; PMID:1607874

A:Accession: P00376

A:Molecule type: genomic RNA

A:Residues: 1-282 <SCH>

C:Genetics:

A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein

C:Keywords: glycoprotein; membrane fusion

Query Match 29.0%; Score 72; DB 2; Length 282;

Best Local Similarity 100.0%; Pred. No. 0.25;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLGV 33

Db 20 LSEIKGVIVHRLGV 34

RESULT 2

P00388

cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)

C:Species: measles virus

C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 24-Nov-1999

C:Accession: P00388

R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.

J. Gen. Virol. 73, 1581-1586, 1992

A>Title: A measles virus isolate from a child with Kawasaki disease: sequence compari

A:Reference number: P00374; MUID:92300360; PMID:1607874

A:Accession: P00388

A:Molecule type: genomic RNA

A:Residues: 1-282 <SCH>

C:Genetics:

A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein

C:Keywords: glycoprotein; membrane fusion

Query Match 29.0%; Score 72; DB 2; Length 282;

Best Local Similarity 100.0%; Pred. No. 0.25;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLGV 33

Db 20 LSEIKGVIVHRLGV 34

```

; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 695
; TYPE: PRT
; ORGANISM: GnRH
US-09-305-924-13

Query Match      38.9%; Score 96.5; DB 11; Length 695;
Best Local Similarity 44.7%; Pred. No. 0.00074;
Matches 21; Conservative 1; Mismatches 10; Indels 15; Gaps 2;

Qy      2 HWSYGLRPGSGGSLKLLSEIKGVIVHRLGVE--GPSLHWSYGLRP 46
        |||||
Db      10 HWSYGLRPGSGSQDWSY-----GLRPGSSQHWSYGLRP 43

Search completed: September 24, 2003, 17:33:10
Job time : 24.5651 secs

```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PMVF-FZD2
US-09-847-102A-33

Query Match          40.7%; Score 101; DB 11; Length 75;
Best Local Similarity 91.3%; Pred. No. 1.5e-05;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 16 LKLLSEIKGVVHRLEGVEGPSL 38
      :||||| ||||||| ||||||| |||
Db 1 MKLLSLIKGVVHRLEGVEGPSL 23

RESULT 12
US-10-285-976-231
; Sequence 231, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Malini, Sen
; APPLICANT: Wu, Christina
; APPLICANT: Leoni, Lorenzo M.
; APPLICANT: Corr, Maripat
; APPLICANT: Carson, Dennis A.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
; FILE REFERENCE: 023070-130320US
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/10/285,976
; PRIOR FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: WO PCT/US02/13802
; PRIOR FILING DATE: 2002-05-01
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 231
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PMVF-2D2
; OTHER INFORMATION: measles virus fusion (MVF) epitope fused to
; OTHER INFORMATION: frizzled domain
US-10-285-976-231

Query Match          40.7%; Score 101; DB 12; Length 75;
Best Local Similarity 91.3%; Pred. No. 1.5e-05;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 16 LKLLSEIKGVVHRLEGVEGPSL 38
      :||||| ||||||| ||||||| |||
Db 1 MKLLSLIKGVVHRLEGVEGPSL 23

RESULT 13
US-09-019-010-4
; Sequence 4, Application US/09019010
; Patent No. US20010014330A1
; GENERAL INFORMATION:
; APPLICANT: HARLAND, RICHARD
; APPLICANT: MANNS, JOHN G.
; APPLICANT: ACRES, STEPHEN D.
; TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
```

```
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,010
; FILING DATE: 05-FEB-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,883
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-019-010-4

Query Match          38.9%; Score 96.5; DB 9; Length 49;
Best Local Similarity 44.7%; Pred. No. 3.4e-05;
Matches 21; Conservative 1; Mismatches 10; Indels 15; Gaps 2;

QY 2 HWSYGLRPGSGPSLKLSEIKGVVHRLEGVE--GPSLHWSYGLRP 46
      :||||| |||
Db 2 HWSYGLRPGSGSQDWSY-----GLRPGSGSQHWSYGLRP 35

RESULT 14
US-09-305-924-11
; Sequence 11, Application US/09305924A
; Publication No. US20030091579A1
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 49
; TYPE: PRT
; ORGANISM: GnRH
US-09-305-924-11

Query Match          38.9%; Score 96.5; DB 11; Length 49;
Best Local Similarity 44.7%; Pred. No. 3.4e-05;
Matches 21; Conservative 1; Mismatches 10; Indels 15; Gaps 2;

QY 2 HWSYGLRPGSGPSLKLSEIKGVVHRLEGVE--GPSLHWSYGLRP 46
      :||||| |||
Db 2 HWSYGLRPGSGSQDWSY-----GLRPGSGSQHWSYGLRP 35

RESULT 15
US-09-305-924-13
; Sequence 13, Application US/09305924A
; Publication No. US20030091579A1
```

; SEQ ID NO 9  
; LENGTH: 45  
; TYPE: PRT  
; ORGANISM: Human  
US-10-076-674-9

Query Match 46.8%; Score 116; DB 12; Length 45;  
Best Local Similarity 85.7%; Pred. No. 1e-07; Indels 2; Gaps 1;  
Matches 24; Conservative 0; Mismatches 2;

Qy 19 LSEIKGVIVHRLGEGPSLHWSYGLRP 46  
|||||  
Db 19 LSEIKGVIVHRLGEGVGE--HWSYGLRP 44  
|||||

## RESULT 8

US-10-223-711-10  
; Sequence 10, Application US/10223711  
; Publication No. US20030113344A1  
; GENERAL INFORMATION:  
; APPLICANT: Bakaletz, Lauren O.  
; APPLICANT: Kaumaya, Pravin T.P.  
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides  
; FILE REFERENCE: 18525/04058  
; CURRENT APPLICATION NUMBER: US/10/223,711  
; CURRENT FILING DATE: 2002-08-19  
; PRIOR APPLICATION NUMBER: 09/148,711  
; PRIOR FILING DATE: 1998-09-04  
; PRIOR APPLICATION NUMBER: 08/460,502  
; PRIOR FILING DATE: 1995-06-02  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-223-711-10

Query Match 42.7%; Score 106; DB 15; Length 40;  
Best Local Similarity 79.3%; Pred. No. 1.7e-06; Indels 0; Gaps 0;  
Matches 23; Conservative 0; Mismatches 6;

Qy 6 GLRPGSGPSLKLSEIKGVIVHRLGVE 34  
|||  
Db 12 GTRDHKKGPSLKLSEIKGVIVHRLGVE 40  
|||||

## RESULT 9

US-09-847-102A-31  
; Sequence 31, Application US/09847102A  
; Publication No. US20030044409A1  
; GENERAL INFORMATION:  
; APPLICANT: University of California  
; APPLICANT: Carson, Dennis A.  
; APPLICANT: Corr, Maripat  
; APPLICANT: Rhee, Chae-Seo  
; APPLICANT: Lorenzo, Leoni M.  
; APPLICANT: Malini, Sen  
; TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR  
; STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS  
; FILE REFERENCE: 22000-20629.00  
; CURRENT APPLICATION NUMBER: US/09/847,102A  
; CURRENT FILING DATE: 2001-05-01  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 31  
; LENGTH: 75  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: PFZD2-MMVF

## US-09-847-102A-31

Query Match 41.1%; Score 102; DB 11; Length 75;  
Best Local Similarity 47.2%; Pred. No. 1.1e-05; Indels 24; Gaps 1;  
Matches 25; Conservative 2; Mismatches 2;

Qy 6 GLRPGSS-----GPSLKLSEIKGVIVHRLGVE 34  
|||  
Db 23 GLOPGAGGTPGGPGGAPPYATLEHPHFCGPSLKLSEIKGVIVHRLGVE 75  
|||||

## RESULT 10

US-10-285-976-229  
; Sequence 229, Application US/10285976  
; Publication No. US20030165500A1  
; GENERAL INFORMATION:  
; APPLICANT: Rhee, Chae-Seo  
; APPLICANT: Malini, Sen  
; APPLICANT: Wu, Christina  
; APPLICANT: Leoni, Lorenzo M.  
; APPLICANT: Corr, Maripat  
; APPLICANT: Carson, Dennis A.  
; TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy  
; TITLE OF INVENTION: in Head and Neck Squamous Cell Carcinomas  
; FILE REFERENCE: 023070-130320US  
; CURRENT APPLICATION NUMBER: US/10/285,976  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: US 60/287,995  
; PRIOR FILING DATE: 2001-05-01  
; PRIOR APPLICATION NUMBER: WO PCT/US02/13802  
; PRIOR FILING DATE: 2002-05-01  
; NUMBER OF SEQ ID NOS: 232  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 229  
; LENGTH: 75  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: PFZD2-MMVF  
; OTHER INFORMATION: measles virus fusion (MVF) epitope fused to  
; OTHER INFORMATION: frizzled domain  
US-10-285-976-229

Query Match 41.1%; Score 102; DB 12; Length 75;  
Best Local Similarity 47.2%; Pred. No. 1.1e-05; Indels 24; Gaps 1;  
Matches 25; Conservative 2; Mismatches 2;

Qy 6 GLRPGSS-----GPSLKLSEIKGVIVHRLGVE 34  
|||  
Db 23 GLOPGAGGTPGGPGGAPPYATLEHPHFCGPSLKLSEIKGVIVHRLGVE 75  
|||||

## RESULT 11

US-09-847-102A-33  
; Sequence 33, Application US/09847102A  
; Publication No. US20030044409A1  
; GENERAL INFORMATION:  
; APPLICANT: University of California  
; APPLICANT: Carson, Dennis A.  
; APPLICANT: Corr, Maripat  
; APPLICANT: Rhee, Chae-Seo  
; APPLICANT: Lorenzo, Leoni M.  
; APPLICANT: Malini, Sen  
; TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR  
; STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS  
; FILE REFERENCE: 22000-20629.00  
; CURRENT APPLICATION NUMBER: US/09/847,102A  
; CURRENT FILING DATE: 2001-05-01  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 33  
; LENGTH: 75

```
; SOFTWARE: PatentIn version 3.1
```

```
; NAME/KEY: PEPTIDE
; LOCATION: (35)..(38)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (39)..(47)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17
```

```
Query Match          99.2%; Score 246; DB 9; Length 47;
Best Local Similarity 100.0%; Pred. No. 3.3e-24;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy  2 HWSYGLRPGSGPSLKLSEIKGVIVHRLEGVEGSLHWSYGLRP 46
      |||||
Db   2 HWSYGLRPGSGPSLKLSEIKGVIVHRLEGVEGSLHWSYGLRP 46
```

## RESULT 2

```
US-09-848-834A-13
; Sequence 13, Application US/09848834A
; Patent No. US20020076416A1
```

## GENERAL INFORMATION:

```
; APPLICANT: Apton Corporation
```

```
; TITLE OF INVENTION: Chimeric Peptide Immunogens
```

```
; FILE REFERENCE: 1102865-0047
```

```
; CURRENT APPLICATION NUMBER: US/09/848,834A
```

```
; CURRENT FILING DATE: 2001-05-04
```

```
; PRIOR APPLICATION NUMBER: 60/202,328
```

```
; PRIOR FILING DATE: 2000-05-05
```

```
; NUMBER OF SEQ ID NOS: 20
```

```
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 13
```

```
; LENGTH: 34
```

```
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
```

## FEATURE:

```
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the GnRH hormone linked by a spacer to amino acid sequence 288-302 of the measles virus fusion protein, F
; OTHER INFORMATION: he Measles virus fusion protein, F
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
```

```
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
```

```
; NAME/KEY: PEPTIDE
```

```
; LOCATION: (11)..(18)
```

```
; OTHER INFORMATION: Spacer peptide
```

```
; NAME/KEY: PEPTIDE
```

```
; LOCATION: (19)..(34)
```

```
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles
```

```
; OTHER INFORMATION: virus fusion protein, F
```

```
; NAME/KEY: MOD_RES
```

```
; LOCATION: (1)..(1)
```

```
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
```

```
US-09-848-834A-13
```

```
Query Match          69.8%; Score 173; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 4.3e-15;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy  2 HWSYGLRPGSGPSLKLSEIKGVIVHRLEGVE 34
      |||||
Db   2 HWSYGLRPGSGPSLKLSEIKGVIVHRLEGVE 34
```

## RESULT 3

```
US-09-848-834A-9
; Sequence 9, Application US/09848834A
; Patent No. US20020076416A1
```

## GENERAL INFORMATION:

```
; APPLICANT: Apton Corporation
```

```
; TITLE OF INVENTION: Chimeric Peptide Immunogens
```

```
; FILE REFERENCE: 1102865-0047
```

```
; CURRENT APPLICATION NUMBER: US/09/848,834A
```

```
; CURRENT FILING DATE: 2001-05-04
```

```
; SEQ ID NO 9
```

```
; LENGTH: 31
```

```
; TYPE: PRT
```

## FEATURE:

```
; ORGANISM: Artificial Sequence
```

```
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of hum
```

```
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of ret
```

```
; OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino ac
```

```
; OTHER INFORMATION: sequence 1-10 of GnRH
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (46)..(46)
; OTHER INFORMATION: Amidated glycine or glycylamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
```

```
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
```

```
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 9
```

```
; LENGTH: 31
```

```
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
```

## FEATURE:

```
; OTHER INFORMATION: Chimeric peptide made up of amino acid sequence 288-302 of the measles virus fusion protein, F linked by a spacer peptide to an
```

```
; OTHER INFORMATION: acid sequence 2-10 of the GnRH hormone
```

```
; NAME/KEY: MOD_RES
```

```
; LOCATION: (1)..(1)
```

```
; OTHER INFORMATION: Amidated Lysine
```

```
; NAME/KEY: PEPTIDE
```

```
; LOCATION: (1)..(15)
```

```
; OTHER INFORMATION: Peptide corresponds to the amino acid sequences 288-302 of the
```

```
; OTHER INFORMATION: measles virus fusion protein, F
```

```
; NAME/KEY: PEPTIDE
```

```
; LOCATION: (19)..(22)
```

```
; OTHER INFORMATION: Spacer peptide
```

```
; NAME/KEY: PEPTIDE
```

```
; LOCATION: (23)..(31)
```

```
; OTHER INFORMATION: Peptide corresponds to amino acid sequences 2-10 of the human
```

```
; OTHER INFORMATION: GnRH hormone
```

```
; NAME/KEY: MOD_RES
```

```
; LOCATION: (31)..(31)
```

```
; OTHER INFORMATION: Amidated glycine or glycylamide
```

```
US-09-848-834A-9
```

```
Query Match          64.1%; Score 159; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy  17 KLLSEIKGVIVHRLEGVEGSLHWSYGLRP 46
      |||||
Db   1 KLLSEIKGVIVHRLEGVEGSLHWSYGLRP 30
```

## RESULT 4

```
US-09-848-834A-19
; Sequence 19, Application US/09848834A
; Patent No. US20020076416A1
```

## GENERAL INFORMATION:

```
; APPLICANT: Apton Corporation
```

```
; TITLE OF INVENTION: Chimeric Peptide Immunogens
```

```
; FILE REFERENCE: 1102865-0047
```

```
; CURRENT APPLICATION NUMBER: US/09/848,834A
```

```
; CURRENT FILING DATE: 2001-05-04
```

```
; PRIOR APPLICATION NUMBER: 60/202,328
```

```
; PRIOR FILING DATE: 2000-05-05
```

```
; NUMBER OF SEQ ID NOS: 20
```

```
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 19
```

```
; LENGTH: 46
```

```
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
```

## FEATURE:

```
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of hum
```

```
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of ret
```

```
; OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino ac
```

```
; OTHER INFORMATION: sequence 1-10 of GnRH
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (46)..(46)
; OTHER INFORMATION: Amidated glycine or glycylamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:57:10 ; Search time 23.5651 Seconds  
(without alignments)  
301.778 Million cell updates/sec

Title: US-09-848-834A-17  
Perfect score: 248  
Sequence: 1 XHWSYGLRSGSPSKLLS.....HRLEGVGPLHWSYGLRPX 47

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues  
Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	246	99.2	47	9	US-09-848-834A-17
2	173	69.8	34	9	US-09-848-834A-13
3	159	64.1	31	9	US-09-848-834A-9
4	147.5	59.5	46	9	US-09-848-834A-19
5	144	58.1	51	9	US-09-848-834A-20
6	141.5	57.1	50	9	US-09-848-834A-18
7	116	46.8	45	12	US-10-076-674-9
8	106	42.7	40	15	US-10-223-711-10
9	102	41.1	75	11	US-09-847-102A-31
10	102	41.1	75	12	US-10-285-976-229
11	101	40.7	75	11	US-09-847-102A-33
12	101	40.7	75	12	US-10-285-976-231
13	96.5	38.9	49	9	US-09-018-010-4
14	96.5	38.9	49	11	US-09-305-924-11
15	96.5	38.9	695	11	US-09-305-924-13

16	96	38.7	40	15	US-10-223-711-11	Sequence 11, Appl
17	94.5	38.1	40	11	US-09-964-201A-35	Sequence 35, Appl
18	94.5	38.1	41	11	US-09-964-201A-34	Sequence 34, Appl
19	89	35.9	31	9	US-09-848-834A-15	Sequence 15, Appl
20	87	35.1	36	9	US-09-848-834A-16	Sequence 16, Appl
21	87	35.1	37	9	US-09-848-834A-14	Sequence 14, Appl
22	86.5	34.9	20	11	US-09-964-201A-26	Sequence 26, Appl
23	86.5	34.9	20	11	US-09-964-201A-29	Sequence 29, Appl
24	86.5	34.9	20	11	US-09-964-201A-30	Sequence 30, Appl
25	86.5	34.9	20	11	US-09-964-201A-31	Sequence 31, Appl
26	79	31.9	34	9	US-09-848-834A-10	Sequence 10, Appl
27	77	31.0	16	9	US-09-848-834A-8	Sequence 8, Appl
28	73	29.4	28	9	US-09-848-834A-11	Sequence 11, Appl
29	73	29.4	33	9	US-09-848-834A-12	Sequence 12, Appl
30	72	29.0	15	11	US-09-747-802-16	Sequence 16, Appl
31	72	29.0	15	11	US-09-747-802-30	Sequence 30, Appl
32	72	29.0	15	11	US-09-865-294-8	Sequence 8, Appl
33	72	29.0	15	11	US-09-865-294-22	Sequence 22, Appl
34	72	29.0	15	12	US-10-261-446-20	Sequence 20, Appl
35	72	29.0	550	10	US-09-873-233A-18	Sequence 18, Appl
36	72	29.0	550	10	US-09-873-233A-20	Sequence 20, Appl
37	71	28.6	17	11	US-09-305-924-4	Sequence 4, Appl
38	69	27.8	15	11	US-09-747-802-37	Sequence 37, Appl
39	69	27.8	15	11	US-09-865-294-29	Sequence 29, Appl
40	69	27.8	16	15	US-10-223-711-3	Sequence 3, Appl
41	69	27.8	19	11	US-09-747-802-48	Sequence 48, Appl
42	69	27.8	19	11	US-09-865-294-40	Sequence 40, Appl
43	66	26.6	19	11	US-09-747-802-51	Sequence 51, Appl
44	66	26.6	19	11	US-09-865-294-43	Sequence 43, Appl
45	66	26.6	662	11	US-09-951-061A-141	Sequence 141, App

ALIGNMENTS

RESULT 1

US-09-848-834A-17  
; Sequence 17, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-03-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17  
; LENGTH: 47  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the  
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of  
; OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid s  
; OTHER INFORMATION: uence 2-10 of the GnRH hormone  
; NAME/KEY: MOD.RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: MOD.RES  
; LOCATION: (47)..(47)  
; OTHER INFORMATION: Amidated-glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(18)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (19)..(34)  
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion prot-ei



Query Match 25.8%; Score 64; DB 1; Length 529;  
 Best Local Similarity 93.3%; Pred. No. 1.4;  
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33  
 DB 291 LSEIKGVIVHRLEGV 305

RESULT 15  
 CONL\_HAPBU  
 ID GONI\_HAPBU STANDARD; PRT; 94 AA.  
 AC P51918; O93387;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)  
 DE (LH-RH I) (Luliberin I).  
 GN GNRH1.  
 OS Haplochromis burtoni (Burton's mouthbrooder).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;  
 OC Cichlidae; Astotilapia.  
 OX NCBI\_TaxID=8153;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=95396797; PubMed=7667296;  
 RX White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;  
 RA "Three gonadotropin-releasing hormone genes in one organism suggest novel roles for an ancient peptide."  
 RT proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99061842; PubMed=9843638;  
 RX White R.B., Fernald R.D.;  
 RA "Ontogeny of gonadotropin-releasing hormone (GnRH) gene expression reveals a distinct origin for GnRH-containing neurons in the midbrain."  
 RT Gen. Comp. Endocrinol. 112:322-329(1998).  
 RN [3]  
 RN SEQUENCE OF 23-32.  
 RC TISSUE=Pituitary;  
 RX MEDLINE=95372591; PubMed=7644702;  
 RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,  
 RA White S.A., Francis R.C., Fernald R.D., Licht P., Warby C.,  
 RA Sherwood N.M.;  
 RT "Primary structure of solitary form of gonadotropin-releasing hormone (GnRH) in cichlid pituitary; three forms of GnRH in brain of cichlid and pumpkinseed fish."  
 RT Regul. Pept. 57:43-53(1995).  
 RL Gen. Comp. Endocrinol. 112:322-329(1995).  
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-GONADAL AXIS.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOTHYSEAL AXONS.  
 CC -!- MASS SPECTROMETRY: MW=1113.9; METHOD=MALDI; RANGE=23-32.  
 CC -!- SIMILARITY: Belongs to the GnRH family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U31865; AAC59691.1; -.  
 DR EMBL; AF076961; AAC27716.1; -.  
 DR FIR; I50739; I50739.  
 DR InterPro; IPR002012; GnRH.  
 DR InterPro; IPR004079; GonadoliberinI.

DR Pfan; PF00446; GnRH; 1.  
 DR PRINTS; PR01541; GONADOLIBRNI.  
 DR PROSITE; PS00473; GnRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal; Multigene family; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 22  
 FT CHAIN 23 94 PROGONADOLIBERIN I.  
 FT PEPTIDE 23 32 GONADOLIBERIN I.  
 FT PEPTIDE 36 94 GnRH-ASSOCIATED PEPTIDE I (POTENTIAL).  
 FT MOD\_RES 23 23 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP).  
 FT CONFLICT 86 94 ENGHRTFKK -> KMDTGHRSNREFL (IN REF. 1).  
 SQ SEQUENCE 94 AA; 10382 MW; E57DBA8333278D7 CRC64;

Query Match 25.6%; Score 63.5; DB 1; Length 94;  
 Best Local Similarity 42.9%; Pred. No. 0.26;  
 Matches 15; Conservative 2; Mismatches 17; Indels 1; Gaps 1;

QY 2 HWSYGLRPGSSGPKLLSEIKGVIVHRLEGVGP 36  
 DB 24 HWSYGLSPGGK-RDLDFNSDTLGNMVEEPPRVEAP 57

Search completed: September 24, 2003, 17:41:21  
 Job time : 9.82299 secs





DR EMBL; M21849; AAA42878.1; -;  
DR EMBL; X65509; CAA46481.1; -;  
DR PIR; JS0321; VGNZCD.  
DR PIR; S21382; S21382.  
DR HSSP; P04849; 1SVF.  
DR InterPro: IPR000776; Fusion\_gly.  
DR Pfam; PF00523; fusion\_gly; 1.  
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
FT SIGNAL 1 ? 662 FUSION GLYCOPROTEIN F0.  
FT CHAIN ? 224 PROTEIN F2.  
FT CHAIN ? 224 PROTEIN F1.  
FT TRANSMEM 606 829 POTENTIAL.  
FT DISULFID 180 307 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
FT CARBOHYD 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 3 3 R -> K (IN REF. 2).  
FT CONFLICT 140 140 D -> N (IN REF. 2).  
FT CONFLICT 152 152 N -> S (IN REF. 2).  
FT CONFLICT 171 171 I -> M (IN REF. 2).  
FT CONFLICT 174 174 A -> V (IN REF. 2).  
FT CONFLICT 662 662 L -> H (IN REF. 2).  
SQ SEQUENCE 662 AA; 72970 MW; FB2C81C9797805F0 CRC64;  
  
Query Match 26.6%; Score 66; DB 1; Length 662;  
Best Local Similarity 73.7%; Pred. No. 1;  
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
QY 15 SLKLLSEIKGVIVHRLGV 33  
DB 396 SYPTLSEIKGVIVHRLGV 414  
  
RESULT 10  
VGLF\_RINDK  
ID VGLF\_RINDK STANDARD; PRT; 546 AA.  
AC P12574;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
DE Fusion glycoprotein F1].  
GN F.  
OS Rinderpest virus (strain Kabete O) (RDV).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
OX NCBI\_TaxID=11242;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88322864; PubMed=3413983;  
RA Hsu D., Yamanaka M., Miller J., Dale B., Grubman M., Yilma T.;  
RT "Cloning of the fusion gene of rinderpest virus: comparative sequence  
RT analysis with other morbilliviruses."  
RL Virology 166:149-153(1988).  
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
CC MEMBRANES.  
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
CC LINKED BY A DISULFIDE BOND.  
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
CC FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M21514; AAA47400.1; -;  
DR PIR; A31051; VGNZK.

DR HSSP; P04849; 1SVF.  
DR InterPro: IPR000776; Fusion\_gly.  
DR Pfam; PF00523; fusion\_gly; 1.  
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.  
FT CHAIN 20 108 F2 PROTEIN.  
FT CHAIN 109 546 F1 PROTEIN.  
FT DOMAIN 104 108 ARG-RICH (BASIC).  
FT TRANSMEM 109 133 POTENTIAL.  
FT TRANSMEM 484 513 POTENTIAL.  
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).  
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 546 AA; 58662 MW; 476D74DCC18BCFCF CRC64;  
  
Query Match 26.2%; Score 65; DB 1; Length 546;  
Best Local Similarity 86.7%; Pred. No. 1.1;  
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 19 LSEIKGVIVHRLGV 33  
DB 284 LSEIKGVIIHLEGV 298  
  
RESULT 11  
VGLF\_PHODV  
ID VGLF\_PHODV STANDARD; PRT; 631 AA.  
AC P28886;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
DE Fusion glycoprotein F1].  
GN F.  
OS Phocine distemper virus (PDV).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
OX NCBI\_TaxID=11240;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Isolate DK88-4A;  
RX MEDLINE=92113538; PubMed=1765768;  
RA Koovanees J., Blixenkron-Moeller M., Sharma B., Oervell C.,  
RA Norby E.;  
RT "The nucleotide sequence and deduced amino acid composition of the  
RT haemagglutinin and fusion proteins of the morbillivirus phocid  
RT distemper virus."  
RL J. Gen. Virol. 72:2959-2966(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Ulster/86;  
RX MEDLINE=92398437; PubMed=1524494;  
RA Curran M.D., Lu Y.J., Rima B.K.;  
RT "The fusion protein gene of phocine distemper virus: nucleotide and  
RT deduced amino acid sequences and a comparison of morbillivirus fusion  
RT proteins."  
RL Arch. Virol. 126:159-169(1992).  
RN [3]  
RP SEQUENCE OF 95-631 FROM N.A.  
RX STRAIN=Ulster/88;  
RX MEDLINE=91089508; PubMed=2264246;  
RA Curran M.D., Loan D.O., Rima B.K., Kennedy S.;  
RT "Nucleotide sequence analysis of phocine distemper virus reveals its  
RT distinctness from canine distemper virus."  
RL Vet. Rec. 127:430-431(1990).  
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
CC MEMBRANES.  
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
CC LINKED BY A DISULFIDE BOND.

```

DR PIR; A28921; VGNZRL.
DR HSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF005523; fusion_gly; 1.
DR Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
KW SIGNAL
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL..
FT TRANSMEM 484 513 POTENTIAL..
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC)..
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL) .
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . ) (POTENTIAL) .
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . ) (POTENTIAL) .
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . ) (POTENTIAL) .
SQ SEQUENCE 546 AA; 58911 MW; 985029418F28FFB5 CRC64;

Query Match 26.6%; Score 66; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.83;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
DB 284 LSEIKGVIVHRLESV 298

RESULT 9
VGLF_CDOVO STANDARD; PRT; 662 AA.
ID ID VGLF_CDOVO
AC PI2569; Q65991;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Canine distemper virus (strain Onderstepoort) (CDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_Taxid=11233;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88129050; PubMed=3433924;
RA Barrett T., Clarke D.K., Evans S.A., Rima B.K.;
RT "The nucleotide sequence of the gene encoding the F protein of canine
RT distemper virus: a comparison of the deduced amino acid sequence with
RT other paramyxoviruses.";
RT Virus Res. 8:373-386(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227696; PubMed=8470428;
RA Wild T.F., Bernard A., Spehner D., Villevall D., Drillien R.;
RT "vaccination of mice against canine distemper virus-induced
RT encephalitis with vaccinia virus recombinants encoding measles or
RT canine distemper virus antigens.";
RL Vaccine 11:438-444(1993)..
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -!- SUBUNIT: THE NATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licens@isb-sib.ch)
CC -----

```

```
RA Billeter M.A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14915; AAA46423.1; -;
DR EMBL; X05597; CAA29090.1; ALT_INIT.
DR EMBL; K01711; AAA75498.1; ALT_INIT.
DR EMBL; K01711; AAA75499.1; -;
DR EMBL; U03657; AAA56647.1; ALT_INIT.
DR EMBL; U03659; AAA56649.1; ALT_INIT.
DR EMBL; U03670; AAA56660.1; ALT_INIT.
DR EMBL; U08416; AAA50550.1; ALT_INIT.
DR EMBL; Z66517; CAA91367.1; ALT_INIT.
DR EMBL; Z66517; CAA91368.1; -;
DR HSSP; P04849; 1SVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.
FT CHAIN 24 112 PROTEIN F2.
FT CHAIN 113 550 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 495 515 POTENTIAL.
FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
FT DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 550 AA; 59532 MW; 7AA4FICA82169093 CRC64;

Query Match 29.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
Db 288 LSEIKGVIVHRLEGV 302

RESULT 6
VGLF_RINDR
ID VGLF_RINDR STANDARD; PRT; 546 AA.
AC P41356;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain RBOK) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=36409;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088609; PubMed=7996154;
RA Evans S.A., Baron M.D., Chamberlain R.W., Coatley L., Barrett T.;
RT "Nucleotide sequence comparisons of the fusion protein gene from
```

```
RT virulent and attenuated strains of rinderpest virus.";
RL J. Gen. Virol. 75:3611-3617(1994).
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z30700; CAA83186.1; -;
DR EMBL; Z30697; CAA83181.1; -;
DR PIR; S47305; S47305.
DR HSSP; P04849; 1SVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58705 MW; ED3DF8AFFDBECB95 CRC64;

Query Match 28.6%; Score 71; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.19;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
Db 284 LSEIKGVIVHRLEGV 298

RESULT 7
GONI_PAGMA
ID GONI_PAGMA STANDARD; PRT; 95 AA.
AC P70074;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I)
DE (LH-RH I) (Luliberin I).
GN GNRH1.
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Pagrus.
OX NCBI_TaxID=143350;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Okuzawa K., Granneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GNRH family.
```



```

QY      19 LSEIKGVIVHRLEGV 33
       |||||||
Db      288 LSEIKGVIVHRLEGV 302

RESULT 3
VGLF_RINDB STANDARD; PRT; 546 AA.
AC PA1360;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2 &
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain RB71) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=39007;
OC [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95088609; PubMed=7996154;
RX Evans S.A., Baron M.D., Chamberslain R.W., Goatley L., Barrett T.
RT "Nucleotide sequence comparisons of the fusion protein gene from
RT virulent and attenuated strains of rinderpest virus.";
RL J. Gen. Virol. 75:3611-3617(1994).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a col-
CC between the Swiss Institute of Bioinformatics and the EMBL out-
CC the European Bioinformatics Institute. There are no restriction
CC use by non-profit institutions as long as its content is in-
CC modified and this statement is not removed. Usage by and for
CC entities requires a license agreement (See http://www.isb-sib.ch)
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z31656; CAAB3482.1; -.
DR PIR; S47300; S47300.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion Protein; Transmembrane; Envelope protein; S
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL.
FT FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL.
FT FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL.
FT FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL.
FT FT CARBOHYD 518 518 O-LINKED (POTENTIAL).
SQ SEQUENCE 546 AA; 58418 MW; 38B539B89344F401 CRC64;

Query Match 29.0%; Score 72; DB 1; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 15; Conservative 0; Mismatches 0; Indels 0;

QY      19 LSEIKGVIVHRLEGV 33
       |||||||
Db      284 LSEIKGVIVHRLEGV 298

```





```
RESULT 14
AAB20151
ID AAB20151 standard; Protein; 109 AA.
XX
AC AAB20151;
XX
DT 30-APR-2001 (first entry)
XX
DE Growth differentiation factor 8 AutoVac construct GDF-8 P30-3B.
XX
KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW cardiant; human; mutant; mutein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Clostridium tetani.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..83
FT /note= "identical to residues 267-349 of human
FT GDF-8"
FT
FT Region 84..104
FT /note= "tetanus toxoid P2 epitope"
FT
FT Region 105..109
FT /note= "identical to residues 371-375 of human
FT GDF-8"
FT
FT Misc-difference 73
FT /note= "Cys-73 may be substituted by Ser to avoid
FT disulfide bond formation"
FT
FT Misc-difference 90..91
FT /note= "optionally replaced by Glu-Gly"
FT
XX
PN WO200105820-A2.
XX
XX 25-JAN-2001.
XX
XX 20-JUL-2000; 2000WO-DK00413.
XX
XX 20-JUL-1999; 99DK-0001014.
XX
XX 26-JUL-1999; 99US-0145275.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Halkier T, Mouritsen S, Klysner S;
XX
XX WPI; 2001-112680/12.
XX
XX Increasing the muscle mass of animals used in meat production by down
XX regulating growth differentiation factor 8 (GDF-8) activity in the
XX animal through induction of anti-GDF-8 antibody production -
XX
XX Example 1; Page 104; 110pp; English.
XX
XX The present sequence is that of AutoVac construct GDF-8 P30-3B,
XX comprising the 109 C-terminal amino acid residues of human
XX growth differentiation factor 8 (GDF-8) in which residues 84-104
XX are replaced by the promiscuous tetanus toxin T-cell epitope P30
XX (see AAB20144). It is an object of the invention to produce a
XX recombinant therapeutic vaccine that is capable of effecting
XX down-regulation of GDF-8 in order to increase the muscle growth
XX rate of farm animals. The vaccines (see AAB20145-53) are capable
XX of breaking autotolerance against autologous GDF-8. They comprise
XX the C-terminal portion of human GDF-8 in which a portion of the
XX native sequence is replaced by a T-cell epitope such as P30, with
XX minimal disturbance of the authentic 3-dimensional structure of
XX the protein. Nucleic acids encoding the GDF-8 variants can be used
XX for genetic immunisation of the animals. Down-regulation of GDF-8
XX activity can increase muscle mass by up to at least 45% in cattle,
XX pigs and poultry used for meat production, reducing the need for
XX antibiotic feed-additives. Anti-GDF8 vaccines can be used to
XX treat human diseases such as cancer cachexia where muscle atrophy is
```

```
CC pronounced and for patients suffering from acute and chronic heart
CC failure.
XX
SQ Sequence 109 AA;
XX
XX Query Match 43.6%; Score 119.5; DB 22; Length 109;
XX Best Local Similarity 67.6%; Pred. No. 4.2e-08;
XX Matches 25; Conservative 1; Mismatches 2; Indels 9; Gaps 1;
XX
Qy 10 GSSGPSL-----FNNETVSPFLRVPKVSASHLE 37
XX |||:| | | | | | | | | | | | | | | | | | | | |
Db 68 GSAGPCTPTKMSPIFNFTVSPFLRVPKVSASHLE 104
XX
RESULT 15
AAY92644
ID AAY92644 standard; Protein; 750 AA.
XX
XX AAY92644;
XX
DT 10-AUG-2000 (first entry)
XX
DE Mutant human prostate specific membrane antigen construct, hPSM6.3.
XX
KW Prostate specific membrane antigen; immunogenized construct; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Peptide 448..462
XX /label= p2
XX /note= "foreign epitope"
XX
XX Peptide 210..230
XX /label= p30
XX /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
XX
XX 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
XX peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page -: 220pp; English.
XX
XX AAY92627-49 are mutant immunogenized human prostate specific membrane
XX antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
XX P30). The immunogenic analogues of PSM can be used in the claimed method
XX as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX antibody binding regions and cysteine residues involved in disulfide
XX bonds are preserved in the immunogenized forms. The method is used for
XX inducing immune responses against weakly immunogenic cell-associated
XX peptide antigens (PA) such as those associated with cancers
XX (self-proteins), e.g. human prostate specific membrane antigen (PSM),
XX heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX method comprises effecting simultaneous presentation by antigen producing
XX cells (APCs) of the animals immune system of: (1) at least 1 CTL
XX (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
```



```
XX PA (MEBI-) M & E BIOTECH AS.
XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G;
XX DR WPI: 2000-349917/30.
XX XX
XX PT Inducing immune responses to weakly immunogenic, tumor associated
XX PT peptide antigens for the treatment of breast and prostate cancer
XX XX
XX PS Example 4; Page -: 220pp; English.
XX CC This is an immunogenized MUC-1 analogue containing foreign epitopes P2
XX CC and P30. Immunogenic analogues of MUC-1 and, e.g. human prostate
XX CC specific membrane antigen (hPSM) can be used in the claimed method as an
XX CC autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX CC antibody binding regions and cysteine residues involved in disulfide
XX CC bonds are preserved in the immunogenized forms (see features table). 10
XX CC regions suitable for the insertion of foreign T helper epitopes were
XX CC identified. The method is used for inducing immune responses against
XX CC weakly immunogenic cell-associated peptide antigens (pA) such as those
XX CC associated with cancers (self-proteins), e.g. hPSM, heregulin 2 (Her2)
XX CC and/or fibroblast growth factor 8b (FGF8b). The method comprises
XX CC effecting simultaneous presentation by antigen producing cells (APCs) of
XX CC the animals immune system of: (1) at least 1 CTL (cytotoxic
XX CC T-lymphocyte) group derived from the PA and/or at least 1 B-cell group
XX CC derived from the cell-associated PA; and (2) at least 1 first T helper
XX CC cell group which is foreign to the animal. Analogues of human PSM, human
XX CC Her2 and human/murine FGF8b comprising a substantial part of all known
XX CC and predicted CTL and B-cell epitopes of the respective PA and including
XX CC at least one foreign T helper epitope are also claimed. The method is
XX CC used to treat prostate, prostate/breast or breast cancer when the PA is
XX CC human PSM, FGF8b and Her2, respectively.
XX CC Note: This sequence does not appear in the specification. It was made
XX CC using the mucin repeat sequence (AA92664), P2 and P30 (AA92625-26),
XX CC which appear on pages 220, 213 and 214 respectively, of the
XX CC specification.
XX SQ Sequence 216 AA;

Query Match 44.2%; Score 121; DB 21; Length 216;
Best Local Similarity 64.1%; Pred. No. 6e-08;
Matches 25; Conservative 2; Mismatches 2; Indels 10; Gaps 1;

QY 9 PGSSGP-----SLFNFTVSFWLRVPKVSASHLE 37
Dbbbbb : bbbbbbbbbb
118 PGSTAPPAHGVTSAPDTRFNFTVSFWLRVPKVSASHLE 156

RESULT 11
AA92627
ID AA92627 standard; Protein: 750 AA.
XX AC AA92627;
XX XX
XX DT 10-AUG-2000 (first entry)
XX DE Mutant human prostate specific membrane antigen construct, hPSM1.1.
XX KW Prostate specific membrane antigen; immunogenized construct; mutant;
XX KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
XX KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX OS Homo sapiens.
XX OS Synthetic.
XX XX
XX FH Key Location/Qualifiers
XX FT Peptide 17..31
XX FT /label= P2
XX FT /note= "foreign epitope"
XX FT Peptide 32..52
XX FT /label= P30
```

```
FT XX /note= "foreign epitope"
XX PN WO200020027-A2.
XX XX 13-APR-2000.
XX PD
XX PF 05-OCT-1999; 99WO-DK00525.
XX XX
XX PR 05-OCT-1998; 98DK-0001261.
XX PR 20-OCT-1998; 98US-0105011.
XX XX
XX PA (MEBI-) M & E BIOTECH AS.
XX XX
XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G;
XX XX
XX DR WPI: 2000-349917/30.
XX XX
XX PT Inducing immune responses to weakly immunogenic, tumor associated
XX PT peptide antigens for the treatment of breast and prostate cancer
XX XX
XX PS Example 1; Page -: 220pp; English.
XX CC AA92627-49 are mutant immunogenized human prostate specific membrane
XX CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
XX CC P30). The immunogenic analogues of PSM can be used in the claimed method
XX CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX CC antibody binding regions and cysteine residues involved in disulfide
XX CC bonds are preserved in the immunogenized forms. The method is used for
XX CC inducing immune responses against weakly immunogenic cell-associated
XX CC peptide antigens (PA) such as those associated with cancers
XX CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
XX CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX CC method comprises effecting simultaneous presentation by antigen producing
XX CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
XX CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
XX CC B-cell group derived from the cell-associated PA; and (2) at least 1
XX CC first T helper cell group which is foreign to the animal. Analogues of
XX CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
XX CC part of all known and predicted CTL and B-cell epitopes of the respective
XX CC PA and including at least one foreign T helper epitope are also claimed.
XX CC The method is used to treat prostate, prostate/breast or breast cancer
XX CC when the PA is human PSM, FGF8b and Her2, respectively.
XX CC Note: This sequence was constructed from the wild type human PSM
XX CC (AA92619), which appears on pages 184-187 of the specification.
XX SQ Sequence 750 AA;

Query Match 44.2%; Score 121; DB 21; Length 750;
Best Local Similarity 95.8%; Pred. No. 2.6e-07;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLEGP 39
Dbbbbb bbbbbbbbbb
31 LFNNFTVSFWLRVPKVSASHLETP 54

RESULT 12
AA92619
ID AA92619 standard; Protein: 109 AA.
XX AC AA92619;
XX XX
XX DT 30-APR-2001 (first entry)
XX DE Growth differentiation factor 8 AutoVac construct GDF-8 P30-2.
XX KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
XX KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
XX KW cardiac; human; mutant; mutein.
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Clostridium tetani.
```



```

SQ Sequence 47 AA;
  Query Match 51.6%; Score 141.5; DB 23; Length 47;
  Best Local Similarity 58.8%; Pred. No. 1.9e-11;
  Matches 30; Conservative 1; Mismatches 11; Indels 9; Gaps 2;

QY 2 HWSYGLRPGSGPSLNFNTVSFWLRVPKVSASHL---EGPSLHWSYGLRP 49
  ||||| ||||| ||||| : | | ||||| |||||
Db 2 HWSYGLRPGSGPSL-----KLLSEIKGVIVHRLEGVGPGLHWSYGLRP 46
  ||||| ||||| ||||| : | | ||||| |||||

RESULT 7
AAU11431
ID AAU11431 standard; peptide; 51 AA.
XX AAU11431;
XX
XX
DT 12-MAR-2002 (first entry)
XX
XX Synthetic immunogen peptide 12.
DE
DE Gonadotropin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
XX Chimeric - Plasmodium falciparum.
OS Chimeric - Mammalia.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 1 /label- OTHER
FT /note= "Other- Pyro-glutamic acid or 5-oxo proline"
FT Peptide 1..10
FT /note= "Gonadotropin releasing hormone epitope"
FT Peptide 11..16
FT /note= "Spacer peptide"
FT Peptide 17..36
FT /note= "Malaria CSP protein (378-398 aa)"
FT Peptide 37..42
FT /note= "Spacer peptide"
FT Peptide 43..51
FT /note= "Gonadotropin releasing hormone epitope"
FT (2-10 aa)"
FT Modified-site 51
FT /note= "Amidated glycine or glycineamide"
XX
XX WO200185763-A2.
XX
XX 15-NOV-2001.
PD
XX 04-MAY-2001; 2001WO-US14363.
XX
XX 05-MAY-2000; 2000US-202328P.
XX
XX (APHT-) APHTON CORP.
XX
XX Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide
PT epitope or its analogue
XX
XX Claim 11; Page 12-13; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH) also known

```

---

```

CC as luteinising hormone releasing hormone, LHRH) comprising a fusion
CC peptide which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic
CC immunogen is useful inducing an immune response against GnRH in an
CC animal subject, and as such is useful as a contraceptive and in the
CC treatment of diseases such as cancer (of the breast, uterus and other
CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic
CC hypertrophy and prostate cancer. The immunogen is effective in eliciting
CC high and specific anti-GnRH antibody titres. The present sequence
CC is a synthetic immunogen of the invention.
XX
XX Sequence 51 AA;
  Query Match 50.5%; Score 138.5; DB 23; Length 51;
  Best Local Similarity 54.7%; Pred. No. 5.1e-11;
  Matches 29; Conservative 4; Mismatches 11; Indels 9; Gaps 2;

QY 2 HWSYGLRPGSGPSLNFNTVSFWLRVPKVSASHL-----SHLEGPSLHWSYGLRP 49
  ||||| ||||| ||||| : | | ||||| |||||
Db 2 HWSYGLRPGSGPSLDEKKIA---KMEKASSVFNVNSSGSPSLHWSYGLRP 50
  ||||| ||||| ||||| : | | ||||| |||||

RESULT 8
AAB20150
ID AAB20150 standard; Protein; 109 AA.
XX
XX AAB20150;
XX
XX 30-APR-2001 (first entry)
XX
XX Growth differentiation factor 8 AutoVac construct GDF-8 P30-3A.
DE
DE Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW cardiant; human; mutant; mutein.
XX
XX Chimeric - Homo sapiens.
OS Chimeric - Clostridium tetani.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Region 1..78
FT /note= "identical to residues 267-345 of human
FT GDF-8"
FT Region 79..99
FT /note= "tetanus toxoid P2 epitope"
FT Region 100..109
FT /note= "identical to residues 366-375 of human
FT GDF-8"
FT Misc-difference 73
FT /note= "Cys-73 may be substituted by Ser to avoid
FT disulfide bond formation"
FT Misc-difference 90..91
FT /note= "optionally replaced by Glu-Gly"
XX
XX WO200105820-A2.
XX
XX 25-JAN-2001.
PD
XX 20-JUL-2000; 2000WO-DK00413.
XX
XX 20-JUL-1999; 99DK-0001014.
XX 26-JUL-1999; 99US-0145275.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Halkier T, Mouritsen S, Klysner S;
XX WPI; 2001-112680/12.
XX
XX Increasing the muscle mass of animals used in meat production by down
PT regulating growth differentiation factor 8 (GDF-8) activity in the
PT animal through induction of anti-GDF-8 antibody production

```

KW luteinising hormone releasing hormone; LHRH; contraceptive;  
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
OS Chimeric - Clostridium tetani.  
OS Chimeric - Mammalia.  
OS Synthetic.  
XX Key Location/Qualifiers  
XX Misc-difference 1 /label= OTHER  
FT Peptide /note= "Other- Pyro-glutamic acid or 5-oxo proline"  
FT 1..10  
FT Peptide /note= "Gonadotrophin releasing hormone epitope"  
FT (1..10 aa)  
FT Peptide /note= "Spacer peptide"  
FT 11..16  
FT Peptide /note= "Tetanus toxoid (830-844 aa)"  
FT 17..31  
FT Peptide /note= "Spacer peptide"  
FT 32..37  
FT Peptide /note= "Spacer peptide"  
FT 38..46  
FT Peptide /note= "Gonadotrophin releasing hormone epitope"  
FT (2-10 aa)  
FT Modified-site 46  
FT /note= "Amidated glycine or glycineamide"  
XX WO200185763-A2.  
XX PN  
XX  
XX  
XX  
PD 15-NOV-2001.  
XX  
XX 04-MAY-2001; 2001WO-US14363.  
XX  
XX 05-MAY-2000; 2000US-202328P.  
XX  
XX (APHT-) APHTON CORP.  
XX  
XX Grimes S, Michaeli D, Stevens VC;  
XX WPI; 2002-049440/06.  
XX  
XX Novel synthetic immunogen for inducing immune response against  
PT gonadotropin releasing hormone, comprises fusion peptide having  
PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
PT epitope or its analogue  
XX  
XX Claim 11; Page 12; 43pp; English.  
XX  
XX The invention relates to a synthetic immunogen for inducing specific  
CC antibodies against gonadotropin releasing hormone (GnRH) also known  
CC as luteinising hormone releasing hormone (LHRH) comprising a fusion  
CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
CC immunomimic peptide epitope or its analogue. The synthetic  
CC immunogen is useful inducing an immune response against GnRH in an  
CC animal subject, and as such is useful as a contraceptive and in the  
CC treatment of diseases such as cancer (of the breast, uterus and other  
CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
CC high and specific anti-GnRH antibody titres. The present sequence  
XX is a synthetic immunogen of the invention.  
SQ Sequence 46 AA;  
Query Match 52.9%; Score 145; DB 23; Length 46;  
Best Local Similarity 60.4%; Pred. No. 6.2e-12;  
Matches 29; Conservative 6; Mismatches 9; Indels 4; Gaps 1;  
QY 2 HWSYGLRPGSGPSLNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 49  
|||||  
Db 2 HWSYGLRPGSGPSLQYIKANSFVIGITELSS----GPSLHWSYGLRP 45  
|||||

RESULT 6  
AAU11428  
ID AAU11428 standard; peptide; 47 AA.  
XX  
AC AAU11428;  
XX  
XX 12-MAR-2002 (first entry)  
XX  
XX Synthetic immunogen peptide 9.  
XX  
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
KW luteinising hormone releasing hormone; LHRH; contraceptive;  
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
XX  
OS Chimeric - Plasmodium falciparum.  
OS Chimeric - Mammalia.  
OS Synthetic.  
XX Key Location/Qualifiers  
XX Misc-difference 1 /label= OTHER  
FT Peptide /note= "Other- Pyro-glutamic acid or 5-oxo proline"  
FT 1..10  
FT Peptide /note= "Gonadotrophin releasing hormone epitope"  
FT (1..10 aa)  
FT Peptide /note= "Spacer peptide"  
FT 11..16  
FT Peptide /note= "Spacer peptide"  
FT 17..34  
FT Peptide /note= "Malaria CSP protein (288-302 aa)"  
FT 35..38  
FT Peptide /note= "Spacer peptide"  
FT 39..47  
FT Peptide /note= "Gonadotrophin releasing hormone epitope"  
FT (2-10 aa)  
FT Modified-site 47  
FT /note= "Amidated glycine or glycineamide"  
XX  
XX WO200185763-A2.  
XX PN  
XX  
XX 15-NOV-2001.  
XX  
XX 04-MAY-2001; 2001WO-US14363.  
XX  
XX 05-MAY-2000; 2000US-202328P.  
XX  
XX (APHT-) APHTON CORP.  
XX  
XX Grimes S, Michaeli D, Stevens VC;  
XX WPI; 2002-049440/06.  
XX  
XX Novel synthetic immunogen for inducing immune response against  
PT gonadotropin releasing hormone, comprises fusion peptide having  
PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
PT epitope or its analogue  
XX  
XX Claim 11; Page 11; 43pp; English.  
XX  
XX The invention relates to a synthetic immunogen for inducing specific  
CC antibodies against gonadotropin releasing hormone (GnRH) also known  
CC as luteinising hormone releasing hormone (LHRH) comprising a fusion  
CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
CC immunomimic peptide epitope or its analogue. The synthetic  
CC immunogen is useful inducing an immune response against GnRH in an  
CC animal subject, and as such is useful as a contraceptive and in the  
CC treatment of diseases such as cancer (of the breast, uterus and other  
CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
CC high and specific anti-GnRH antibody titres. The present sequence  
XX is a synthetic immunogen of the invention.  
XX

OS Synthetic.  
 XX Key Location/Qualifiers  
 FH Peptide 1..21 "Tetanus toxoid sequence (947-967 aa)"  
 FT Peptide /note= "Spacer peptide"  
 FT Peptide 22..25  
 FT Peptide /note= "Gonadotrophin releasing hormone epitope"  
 FT Modified-site 34  
 FT /note= "Amidated glycine or glycinamide"  
 XX WO200185763-A2.  
 PN 15-NOV-2001.  
 PD 04-MAY-2001; 2001WO-US14363.  
 XX 05-MAY-2000; 2000US-202328P.  
 PR (APHT-) APHTON CORP.  
 XX Grimes S, Michaeli D, Stevens VC;  
 PI WPI; 2002-049440/06.  
 XX Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue  
 XX Claim 11; Page 7; 43pp; English.  
 XX The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinizing hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX SQ Sequence 34 AA;  
 Query Match 67.5%; Score 185; DB 23; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 2.le-17;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 17 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 49  
 DB 1 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 33  
 RESULT 4  
 AAR62702  
 ID AAR62702 standard; peptide; 32 AA.  
 XX AAR62702;  
 AC AAR62702;  
 XX 25-MAR-2003 (updated)  
 DT 10-SEP-1995 (first entry)  
 XX LHRH-containing immunogenic peptide.  
 XX Helper T cell epitope; universal immune stimulator; invasin; haptens;  
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;  
 KW androgen-dependent carcinoma; antitumour; infertility;  
 KW tetanus toxin.  
 XX

OS Synthetic.  
 XX Key Location/Qualifiers  
 FH Domain 1..22 /note= "tetanus toxin helper T cell epitope"  
 FT Domain 23..32 /note= "LHRH haptens"  
 XX WO9425060-A1.  
 PN 10-NOV-1994.  
 PD 28-APR-1994; 94WO-US04832.  
 XX 27-APR-1993; 93US-0057166.  
 PR 14-APR-1994; 94US-0229275.  
 XX (LADD/) LADD A E.  
 PA (WANG/) WANG C Y.  
 PA (ZAMB/) ZAMB T.  
 XX Ladd AE, Wang CY, Zamb T;  
 PI WPI; 1994-357910/44.  
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -  
 PT that suppress LHRH activity in males and females  
 XX Claim 8; Page 84; 213pp; English.  
 XX Synthetic immunogenic peptides are provided in which a universal immune  
 CC stimulator is linked to a peptide or protein haptens containing B cell  
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes  
 CC potent immune responses to the coupled peptide or protein. The  
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)  
 CC which elicits an immune response to the coupled peptide in members of  
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)  
 CC an adjuvant peptide sequence from the invasin protein of Yersinia.  
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the  
 CC invasin and Th domains and between the immune stimulator and haptens  
 CC components. When the haptens is LHRH, then optionally the invasin domain  
 CC can be omitted from the immune stimulator component.  
 CC The present sequence represents an LHRH-containing, invasin-free  
 CC immunogenic peptide as above which can be used as a potent vaccine for  
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,  
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign  
 CC uterine tumours, recurrent functional ovarian cysts, (severe)  
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for  
 CC induction of infertility.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX SQ Sequence 32 AA;  
 Query Match 55.5%; Score 152; DB 15; Length 32;  
 Best Local Similarity 87.9%; Pred. No. 4.7e-13;  
 Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;  
 QY 17 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 49  
 DB 3 FNNFTVFWLRVPKVSASHLE---HWSYGLRP 31  
 RESULT 5  
 AAU11430  
 ID AAU11430 standard; peptide; 46 AA.  
 XX AAU11430;  
 AC AAU11430;  
 XX 12-MAR-2002 (first entry)  
 DT Synthetic immunogen peptide 11.  
 DE Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
 KW

FT Peptide /note= "Spacer peptide"  
FT 42..50  
FT /note= "Gonadotrophin releasing hormone epitope  
FT (2-10 aa)"  
FT Modified-site  
FT 50  
FT /note= "Amidated glycine or glycineamide"  
PN WO200185763-A2.  
XX 15-NOV-2001.  
XX  
XX  
XX 04-MAY-2001; 2001WO-US14363.  
XX  
XX 05-MAY-2000; 2000US-202328P.  
XX (APHT-) APHTON CORP.  
XX Grimes S, Michaeli D, Stevens VC;  
XX WPI; 2002-049440/06.  
XX Novel synthetic immunogen for inducing immune response against  
XX gonadotropin releasing hormone, comprises fusion peptide having  
XX promiscuous helper T-cell peptide epitope and immunomimic peptide  
XX epitope or its analogue  
XX Claim 11; Page 11; 43pp; English.  
XX The invention relates to a synthetic immunogen for inducing specific  
XX antibodies against gonadotropin releasing hormone (GnRH) also known  
XX as luteinising hormone releasing hormone, LHRR) comprising a fusion  
XX peptide which comprises a promiscuous helper T-cell peptide epitope and  
XX immunomimic peptide epitope or its analogue. The synthetic  
XX immunogen is useful inducing an immune response against GnRH in an  
XX animal subject, and as such is useful as a contraceptive and in the  
XX treatment of diseases such as cancer (of the breast, uterus and other  
XX gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
XX hypertrophy and prostate cancer. The immunogen is effective in eliciting  
XX high and specific anti-GnRH antibody titres. The present sequence  
XX is a synthetic immunogen of the invention.  
SQ Sequence 50 AA;  
Query Match 99.3%; Score 272; DB 23; Length 50;  
Best Local Similarity 100.0%; Pred. No. 8.6e-29;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 HWSYGLRPGSGPSLFNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 49  
Db ||||| 2 HWSYGLRPGSGPSLFNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 49  
RESULT 2  
AAU11425  
ID AAU11425 standard; peptide; 37 AA.  
XX  
XX AAU11425;  
XX  
XX 12-MAR-2002 (first entry)  
XX  
XX Synthetic immunogen peptide 6.  
XX Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
XX luteinising hormone releasing hormone; LHRR; contraceptive;  
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
XX Chimeric - Clostridium tetani.  
OS Chimeric - Mammalia.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
PH

Misc-difference 1 /label= OTHER  
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
FT 1..10  
FT Peptide /note= "Gonadotrophin releasing hormone epitope"  
FT 11..16  
FT Peptide /note= "Spacer peptide"  
FT 17..37  
FT Peptide /note= "Tetanus toxoid sequence (947-967 aa)"  
XX  
XX WO200185763-A2.  
PN  
XX  
XX 15-NOV-2001.  
XX  
XX 04-MAY-2001; 2001WO-US14363.  
XX  
XX 05-MAY-2000; 2000US-202328P.  
XX (APHT-) APHTON CORP.  
XX Grimes S, Michaeli D, Stevens VC;  
XX WPI; 2002-049440/06.  
XX Novel synthetic immunogen for inducing immune response against  
XX gonadotropin releasing hormone, comprises fusion peptide having  
XX promiscuous helper T-cell peptide epitope and immunomimic peptide  
XX epitope or its analogue  
XX Claim 11; Page 9; 43pp; English.  
XX The invention relates to a synthetic immunogen for inducing specific  
XX antibodies against gonadotropin releasing hormone (GnRH) also known  
XX as luteinising hormone releasing hormone, LHRR) comprising a fusion  
XX peptide which comprises a promiscuous helper T-cell peptide epitope and  
XX immunomimic peptide epitope or its analogue. The synthetic  
XX immunogen is useful inducing an immune response against GnRH in an  
XX animal subject, and as such is useful as a contraceptive and in the  
XX treatment of diseases such as cancer (of the breast, uterus and other  
XX gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
XX hypertrophy and prostate cancer. The immunogen is effective in eliciting  
XX high and specific anti-GnRH antibody titres. The present sequence  
XX is a synthetic immunogen of the invention.  
SQ Sequence 37 AA;  
Query Match 72.6%; Score 199; DB 23; Length 37;  
Best Local Similarity 100.0%; Pred. No. 3.1e-19;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 HWSYGLRPGSGPSLFNNFTVSFWLRVPKVSASHLE 37  
Db ||||| 2 HWSYGLRPGSGPSLFNNFTVSFWLRVPKVSASHLE 37  
RESULT 3  
AAU11421  
ID AAU11421 standard; peptide; 34 AA.  
XX  
XX AAU11421;  
XX  
XX 12-MAR-2002 (first entry)  
XX  
XX Synthetic immunogen peptide 2.  
XX Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
XX luteinising hormone releasing hormone; LHRR; contraceptive;  
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
XX Chimeric - Clostridium tetani.  
OS Chimeric - Mammalia.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:58:36 ; Search time 56.5097 Seconds  
(without alignments)  
140.442 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHWSYGLRPGSGSGLFNF.....VSASHLEGLSHWSYGLRPX 50

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	272	99.3	50	AAU11429	Synthetic immunoge
2	199	72.6	37	AAU11425	Synthetic immunoge
3	185	67.5	34	AAU11421	Synthetic immunoge
4	152	55.5	32	AAU11420	LHRH-containing im
5	145	52.9	46	AAU11430	Synthetic immunoge
6	141.5	51.6	47	AAU11428	Synthetic immunoge
7	138.5	50.5	51	AAU11431	Synthetic immunoge
8	123	44.9	109	AAU11431	Growth differentia
9	121.5	44.3	750	AAU11431	Mutant human prost

10	121	44.2	216	21	AAU11429	MUC-1 analogue con
11	121	44.2	750	21	AAU11429	Mutant human prost
12	120.5	44.0	109	22	AAU11429	Growth differentia
13	120	43.8	188	21	AAU11429	An osteoprotegerin
14	119.5	43.6	109	22	AAU11429	Growth differentia
15	119.5	43.6	750	21	AAU11429	Mutant human prost
16	119	43.4	122	21	AAU11429	Modified murine in
17	119	43.4	122	21	AAU11429	Modified murine in
18	118.5	43.2	31	21	AAU11429	PSNep012 - p30 in
19	118.5	43.2	158	19	AAU11429	TNF30-1, a TNF- $\alpha$ p
20	118.5	43.2	158	23	AAU11429	Human TNF- $\alpha$ an
21	118.5	43.2	693	21	AAU11429	Mutant human PSM a
22	118.5	43.2	693	21	AAU11429	Mutant human PSM a
23	118.5	43.2	750	21	AAU11429	Mutant human prost
24	118.5	43.2	750	21	AAU11429	Mutant human prost
25	118	43.1	158	19	AAU11429	TNF30-5, a TNF- $\alpha$ p
26	118	43.1	158	23	AAU11429	Human TNF- $\alpha$ an
27	118	43.1	703	21	AAU11429	Mutant murine PSM
28	118	43.1	761	21	AAU11429	Mutant murine pros
29	117	42.7	124	21	AAU11429	Modified human int
30	117	42.7	124	21	AAU11429	Modified human int
31	117	42.7	160	22	AAU11429	Growth differentia
32	117	42.7	750	21	AAU11429	Mutant human prost
33	116	42.3	31	21	AAU11429	PSNep010 - p30 in
34	116	42.3	43	22	AAU11429	Tetanus toxoid 830
35	116	42.3	43	22	AAU11429	Amyloid beta/tetan
36	116	42.3	72	22	AAU11429	Tetanus toxoid epi
37	116	42.3	109	22	AAU11429	Growth differentia
38	116	42.3	136	22	AAU11429	Amyloid beta tetan
39	116	42.3	145	21	AAU11429	Modified murine in
40	116	42.3	147	21	AAU11429	Modified human int
41	116	42.3	158	19	AAU11429	TNF30-3, a TNF- $\alpha$ p
42	116	42.3	158	23	AAU11429	Human TNF- $\alpha$ an
43	116	42.3	254	22	AAU11429	Growth differentia
44	116	42.3	750	21	AAU11429	Mutant human prost
45	116	42.3	750	21	AAU11429	Mutant human prost

# ALIGNMENTS

RESULT 1

AAU11429

ID AAU11429 standard; peptide; 50 AA.

XX AAU11429;

AC AAU11429;

DT 12-MAR-2002 (first entry)

DE Synthetic immunogen peptide 10.

XX Gonadotrophin releasing hormone; GnRH; synthetic immunogen;

KW luteinising hormone releasing hormone; LHRH; contraceptive;

KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;

KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;

KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Chimeric - Clostridium tetani.

OS Chimeric - Mammalia.

OS Synthetic.

FT Key Location/Qualifiers

FT Misc-difference 1

FT /label= OTHER

FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"

FT Peptide 1..10

FT /note= "Gonadotrophin releasing hormone epitope

FT (1..10 aa)"

FT Peptide 11..16

FT /note= "Spacer peptide"

FT Peptide 17..37

FT /note= "Tetanus toxoid (947-967 aa)"

FT Peptide 38..41



Db 288 LSEIKGVIVHRLGV 302  
|||||

RESULT 13  
Q9VJ94 PRELIMINARY; PRT; 550 AA.  
AC Q9VJ94; 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Fusion protein.  
GN F.  
OS Measles virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11234;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-G954;  
RX MEDLINE=21635526; PubMed=11773423;  
RX Waku Koumou D., Wild T.F.;  
RT "Adaptation of wild-type measles virus to tissue culture.";  
RL J. Virol. 76:1505-1509(2002).  
DR EMBL; AF059392; AAL29688.1; -.  
DR InterPro: IPR000776; Fusion\_gly.  
DR Pfam: PF00523; fusion\_gly; 1.  
SQ SEQUENCE 550 AA; 59551 MW; 9A7A4BA99B4DA8E9 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33  
Db 288 LSEIKGVIVHRLGV 302  
|||||

RESULT 14  
Q9VJ94 PRELIMINARY; PRT; 550 AA.  
AC Q9VJ94; 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE Fusion protein.  
OS Measles virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11234;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-9301V;  
RX MEDLINE=98440529; PubMed=9765410;  
RX Takeda M., Kato A., Kobune F., Sakata H., Li Y., Shioda T., Sakai Y.,  
RA Asakawa M., Nagai Y.;  
RT "Measles virus attenuation associated with transcriptional impediment  
and a few amino acid changes in the polymerase and accessory  
proteins.";  
RL J. Virol. 72:8690-8696(1998).  
DR EMBL; AB012949; BAA33877.1; -.  
DR EMBL; AB012948; BAA33871.1; -.  
DR HSSP; P04849; 1SVF.  
DR InterPro: IPR000776; Fusion\_gly.  
DR Pfam: PF00523; fusion\_gly; 1.  
SQ SEQUENCE 550 AA; 59512 MW; 7AA4F1D117197BF9 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33  
|||||

Db 288 LSEIKGVIVHRLGV 302

RESULT 15  
Q9QEX1 PRELIMINARY; PRT; 550 AA.  
AC Q9QEX1;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Fusion protein.  
OS Measles virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11234;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Masusako;  
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,  
RA Oqura H.;  
RT "Nucleotide sequences of the fusion protein gene of subacute  
sclerosing panencephalitis viruses: deduced amino acid sequences  
showed the cytoplasmic domain highly mutated --truncated, elongated or  
predicted secondary structure changed.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF179430; AAF02695.1; -.  
DR HSSP; P04849; 1SVF.  
DR InterPro: IPR000776; Fusion\_gly.  
DR Pfam: PF00523; fusion\_gly; 1.  
SQ SEQUENCE 550 AA; 59559 MW; 609EE024A7E59C54 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33  
Db 288 LSEIKGVIVHRLGV 302  
|||||

Search completed: September 24, 2003, 17:47:22  
Job time : 46.5679 secs

P90330  
 ID P90330 PRELIMINARY; PRT; 550 AA.  
 AC P90330;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Nagahata(HB);  
 RA Sheng J., Watanabe M., Ueda S.;  
 RT "Selection of a neurotropic variant of measles virus.";  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Nagahata(HB);  
 RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;  
 RT "An amino acid alteration of F protein responsible for the enhanced  
 RT fusogenicity of measles virus.";  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; D63924; BAA09951.1; -;  
 DR HSSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59589 MW; 73E7BD457ABA39B7 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33  
 |||||  
 DB 288 LSEIKGVIVHRLGV 302

## RESULT 10

Q9QEW7  
 ID Q9QEW7 PRELIMINARY; PRT; 550 AA.  
 AC Q9QEW7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OSA-2;  
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,  
 RA Ogura H.;  
 RT "Nucleotide sequences of the fusion protein gene of subacute  
 RT sclerosing panencephalitis viruses: deduced amino acid sequences  
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or  
 RT predicted secondary structure changed.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF179438; AAF02703.1; -;  
 DR HSSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59333 MW; 086E51FED5582BBA CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33

Db 288 LSEIKGVIVHRLGV 302  
 |||||

## RESULT 11

Q9WMK4  
 ID Q9WMK4 PRELIMINARY; PRT; 550 AA.  
 AC Q9WMK4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WTF;  
 RX MEDLINE=99329215; PubMed=10400788;  
 RA Johnston I.C., Ter Meulen V., Schneider-Schaulies J.,  
 RA Schneider-Schaulies S.;  
 RT "A recombinant measles vaccine virus expressing wild-type  
 RT glycoproteins : consequences for viral spread and cell tropism.";  
 RL J. Virol. 73:6903-6915(1999).  
 DR EMBL; AJ133108; CAB38075.1; -;  
 DR HSSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59580 MW; 8255499968B5D862 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33  
 |||||  
 DB 288 LSEIKGVIVHRLGV 302

## RESULT 12

Q89495  
 ID Q89495 PRELIMINARY; PRT; 550 AA.  
 AC Q89495;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92230209; PubMed=1566568;  
 RA Rota J.S., Hummel K.B., Rota P.A., Bellini W.J.;  
 RT "Genetic variability of the glycoprotein genes of current wild-type  
 RT measles isolates.";  
 RL Virology 188:135-142(1992).  
 DR EMBL; M81903; AAA46422.1; -;  
 DR EMBL; M81901; AAA46421.1; -;  
 DR HSSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59564 MW; A78EC9CD6268E58 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33

```

RP SEQUENCE FROM N.A.
RC STRAIN=K;
RA AYANOT P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA Gusev A.A.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035887; AAK63190.1; -.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 546 AA; 58572 MW; 449B2B2DD7405F0B CRC64;

Query Match          29.0%; Score 72; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
Db 284 LSEIKGVIVHRLGV 298

RESULT 6
P90331 PRELIMINARY; PRT; 550 AA.
AC P90331;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OC NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGAHATA;
RA Sheng J., Watanabe M., Ueda S.;
RT "Selection of a neurotropic variant of measles virus.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGAHATA;
RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT "An amino acid alteration of F protein responsible for the enhanced
RT fusogenicity of measles virus.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGAHATA;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; D63926; BAA09958.1; -.
DR EMBL; AF179431; AAF02696.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59530 MW; 97C991C7E2169839 CRC64;

Query Match          29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
Db 288 LSEIKGVIVHRLGV 302

RESULT 7
Q9QEX0 PRELIMINARY; PRT; 550 AA.
AC Q9QEX0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OC NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Toyoshima;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179432; AAF02697.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59504 MW; 2AA969D37FA5CA17 CRC64;

Query Match          29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
Db 288 LSEIKGVIVHRLGV 302

RESULT 8
Q9QEW9 PRELIMINARY; PRT; 550 AA.
AC Q9QEW9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Fusion protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OC NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179436; AAF02701.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59405 MW; 0AF6DBFC5DD22BBA CRC64;

Query Match          29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
Db 288 LSEIKGVIVHRLGV 302

RESULT 9

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:29:11; Search time 45.5679 Seconds  
(without alignments)  
266.163 Million cell updates/sec

Title: US-09-848-834A-17  
Perfect score: 248  
Sequence: 1 XHWSYGLRPGSSGSPSLKLS.....HRLEGVEGSLHWSYGLRXP 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	72	5	29.2	87	13	Q9YI26
2	72	5	29.0	534	12	Q04243
3	72	5	29.0	537	12	Q04242
4	72	5	29.0	545	12	Q9PXA4
5	72	5	29.0	546	12	Q9LHA5
6	72	5	29.0	550	12	P90331
7	72	5	29.0	550	12	Q9QEX0
8	72	5	29.0	550	12	Q9QEW9
9	72	5	29.0	550	12	P90330
10	72	5	29.0	550	12	Q9QEW7
11	72	5	29.0	550	12	Q9WMK4
12	72	5	29.0	550	12	Q89495
13	72	5	29.0	550	12	Q8V049
14	72	5	29.0	550	12	Q9YJ94
15	72	5	29.0	550	12	Q9QEX1
16	72	5	29.0	550	12	Q9QEW8

17	72	29.0	553	12	Q93055	measles vir
18	72	29.0	553	12	Q9IC36	measles vir
19	72	29.0	553	12	P88973	measles vir
20	72	29.0	553	12	Q83536	measles vir
21	72	29.0	553	12	O11383	measles vir
22	72	29.0	553	12	Q9IFK2	measles vir
23	72	29.0	553	12	Q83533	measles vir
24	72	29.0	553	12	Q83525	measles vir
25	72	29.0	553	12	Q83518	measles vir
26	72	29.0	553	12	P88974	measles vir
27	72	29.0	553	12	Q83527	measles vir
28	72	29.0	553	12	Q83521	measles vir
29	72	29.0	553	12	Q83530	measles vir
30	72	29.0	553	12	Q91248	measles vir
31	72	29.0	553	12	Q91QP2	measles vir
32	72	29.0	553	12	Q04244	measles vir
33	72	29.0	579	12	Q9PW04	measles vir
34	71.5	28.8	68	13	Q8JIF4	measles vir
35	68.5	27.6	64	13	Q8JIF3	measles vir
36	68.5	27.6	64	13	Q8JIF2	measles vir
37	68	27.4	545	12	Q9QEW6	measles vir
38	68	27.4	553	12	O11380	measles vir
39	67.5	27.2	96	13	Q8UW80	measles vir
40	66.5	26.8	552	12	Q66147	measles vir
41	66	26.6	528	12	Q9YJW9	measles vir
42	66	26.6	530	12	Q8QV06	measles vir
43	66	26.6	662	12	Q9DX22	measles vir
44	66	26.6	662	12	Q91KN3	measles vir
45	66	26.6	662	12	Q9YKL7	measles vir

ALIGNMENTS

RESULT 1

Q9YI26	PRELIMINARY;	PRT;	87	AA.
ID	Q9YI26			
AC	Q9YI26;			
DT	01-MAY-1999 (Tremblrel. 10, Created)			
DT	01-MAY-1999 (Tremblrel. 10, Last sequence update)			
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)			
DE	SbGnRH (Gonadoliberin) (Gonadotropin-releasing hormone) (LH-RH)			
DE	(Luliberin) (Fragment).			
OS	Sparus aurata (Gilthead sea bream).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;			
OC	Sparidae; Sparus.			
OX	NCBI_TaxID=8175;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Ovary;			
RA	Nabissi M.;			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
CC	!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY SIMILARITY).			
CC	!- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
DR	EMBL; AF046801; AAD02427.1; -.			
DR	InterPro; IPR002012; GnRH.			
DR	InterPro; IPR004079; Gonadoliberin1.			
DR	Pfam; PF00446; GnRH; 1.			
DR	PRINTS; PR01541; GONADOLIBRNI.			
DR	PROSITE; PS00473; GNRH; 1.			
KW	Amidation; Hormone.			
FT	NON_TER 1			
FT	NON_TER 87			
SQ	SEQUENCE 87 AA; 9871 MW; 0D2463533D96782A CRC64;			
Query Match	29.2%;	Score	72.5;	DB 13; Length 87;
Best Local Similarity	45.7%;	Pred. No.	0.12;	Indels 1; Gaps 1;
Matches	16;	Conservative	3;	Mismatches 15;
OY	2	HWSYGLRPGSSGSPSLKLSLKLVHVRLEGVEGP	36	

RESULT 14  
US-08-913-880C-13  
; Sequence 13, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913,880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 13  
; LENGTH: 865  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 451 to 1315 of SEQ ID NO: 1  
US-08-913-880C-13

Query Match 41.6%; Score 114; DB 4; Length 865;  
Best Local Similarity 95.5%; Pred. No. 5.4e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 16 LFNNFTVSEWLRVPKVSASHLE 37  
Db 496 MFNNFTVSEWLRVPKVSASHLE 517

RESULT 15  
US-08-913-880C-12  
; Sequence 12, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913,880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 12  
; LENGTH: 866  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 450 to 1315 of SEQ ID NO: 1  
US-08-913-880C-12

Query Match 41.6%; Score 114; DB 4; Length 866;  
Best Local Similarity 95.5%; Pred. No. 5.4e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 16 LFNNFTVSEWLRVPKVSASHLE 37  
Db 497 MFNNFTVSEWLRVPKVSASHLE 518

Search completed: September 24, 2003, 17:00:25  
Job time : 14.9199 secs

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/668,381A
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,473
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,164
; REFERENCE/DOCKET NUMBER: 00786/269001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-668-381A-5

```

```

Query Match 41.6%; Score 114; DB 1; Length 618;
Best Local Similarity 95.5%; Pred. No. 3.5e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 16 LFNNFTVSFWLRVPKVSASHLE 37
:|||||
DB 249 MFNNFTVSFWLRVPKVSASHLE 270

```

```

RESULT 10
US-08-913-880C-17
; Sequence 17, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 17
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 463 to 1315 of SEQ ID NO: 1
US-08-913-880C-17

```

```

Query Match 41.6%; Score 114; DB 4; Length 853;
Best Local Similarity 95.5%; Pred. No. 5.3e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 16 LFNNFTVSFWLRVPKVSASHLE 37
:|||||
DB 484 MFNNFTVSFWLRVPKVSASHLE 505

```

```

RESULT 11
US-08-913-880C-16
; Sequence 16, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17

```

```

; SEQ ID NO 16
; LENGTH: 858
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 458 to 1315 of SEQ ID NO: 1
US-08-913-880C-16

```

```

Query Match 41.6%; Score 114; DB 4; Length 858;
Best Local Similarity 95.5%; Pred. No. 5.3e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 16 LFNNFTVSFWLRVPKVSASHLE 37
:|||||
DB 489 MFNNFTVSFWLRVPKVSASHLE 510

```

```

RESULT 12
US-08-913-880C-15
; Sequence 15, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 15
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 456 to 1315 of SEQ ID NO: 1
US-08-913-880C-15

```

```

Query Match 41.6%; Score 114; DB 4; Length 860;
Best Local Similarity 95.5%; Pred. No. 5.3e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 16 LFNNFTVSFWLRVPKVSASHLE 37
:|||||
DB 491 MFNNFTVSFWLRVPKVSASHLE 512

```

```

RESULT 13
US-08-913-880C-14
; Sequence 14, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 14
; LENGTH: 862
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 454 to 1315 of SEQ ID NO: 1
US-08-913-880C-14

```

```

Query Match 41.6%; Score 114; DB 4; Length 862;
Best Local Similarity 95.5%; Pred. No. 5.3e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 16 LFNNFTVSFWLRVPKVSASHLE 37
:|||||
DB 493 MFNNFTVSFWLRVPKVSASHLE 514

```

; Sequence 2, Application US/08280228  
; Patent No. 5571694  
; GENERAL INFORMATION:  
; APPLICANT: Makoff Dr, Andrew J  
; APPLICANT: Romanos Dr, Michael A  
; APPLICANT: Clare Dr, Jeffrey J  
; APPLICANT: Fairweather Dr, Neil F  
; TITLE OF INVENTION: VACCINES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 No. 5571694th Glebe Road  
; CITY: Arlington,  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,228  
; FILING DATE: 25-JUL-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8926832.0  
; FILING DATE: 27-NOV-1990  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9006097.1  
; FILING DATE: 17-MAR-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilson, Mary J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 117-163  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; TELEX: 200797 NIXN UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 452 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-280-228-2

Query Match 41.6%; Score 114; DB 1; Length 452;  
Best Local Similarity 95.5%; Pred. No. 2.4e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRPKVSASHLE 37  
:|||||  
Db 83 MFNNFTVSFWLRPKVSASHLE 104

RESULT 8  
US-08-280-228-4  
; Sequence 4, Application US/08280228  
; Patent No. 5571694  
; GENERAL INFORMATION:  
; APPLICANT: Makoff Dr, Andrew J  
; APPLICANT: Romanos Dr, Michael A  
; APPLICANT: Clare Dr, Jeffrey J  
; APPLICANT: Fairweather Dr, Neil F  
; TITLE OF INVENTION: VACCINES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.

; STREET: 1100 No. 5571694th Glebe Road  
; CITY: Arlington,  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,228  
; FILING DATE: 25-JUL-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/618,312  
; FILING DATE: 27-NOV-1990  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8926832.0  
; FILING DATE: 28-NOV-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9006097.1  
; FILING DATE: 17-MAR-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilson, Mary J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 117-163  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; TELEX: 200797 NIXN UR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 452 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-280-228-4

Query Match 41.6%; Score 114; DB 1; Length 452;  
Best Local Similarity 95.5%; Pred. No. 2.4e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRPKVSASHLE 37  
:|||||  
Db 83 MFNNFTVSFWLRPKVSASHLE 104

RESULT 9  
US-08-668-381A-5  
; Sequence 5, Application US/08668381A  
; Patent No. 5780024  
; GENERAL INFORMATION:  
; APPLICANT: Brown, Robert H.  
; APPLICANT: Fishman, Paul S.  
; APPLICANT: Francis, Jonathan W.  
; APPLICANT: Hosler, Betsy A.  
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN  
; TITLE OF INVENTION: FRAGMENT C HYBRID PROTEIN  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

APPLICATION NUMBER: GB 9006097.1  
FILING DATE: 17-MAR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford Mr, Arthur R  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 510-51  
TELEPHONE: 0101 703 8750400  
TELEFAX: 0101 703 5253468  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-618-312A-2

Query Match 41.6%; Score 114; DB 1; Length 452;  
Best Local Similarity 95.5%; Pred. No. 2.4e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRPKVSASHLE 37  
Db 83 MFNNFTVSFWLRPKVSASHLE 104

RESULT 5  
US-07-618-312A-4  
Sequence 4, Application US/07618312A  
Patent No. 5389540  
GENERAL INFORMATION:  
APPLICANT: Makoff Dr, Andrew J  
APPLICANT: Romanos Dr, Michael A  
APPLICANT: Clare Dr, Jeffrey J  
APPLICANT: Fairweather Dr, Neil F  
TITLE OF INVENTION: VACCINES  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: 14th Floor  
STREET: 2200 Clarendon Boulevard,  
CITY: Arlington,  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/618,312A  
FILING DATE: 19910516  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8926832.0  
FILING DATE: 28-NOV-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9006097.1  
FILING DATE: 17-MAR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford Mr, Arthur R  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 510-51  
TELEPHONE: 0101 703 8750400  
TELEFAX: 0101 703 5253468  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-07-618-312A-4  
Query Match 41.6%; Score 114; DB 1; Length 452;  
Best Local Similarity 95.5%; Pred. No. 2.4e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 16 LFNNFTVSFWLRPKVSASHLE 37  
Db 83 MFNNFTVSFWLRPKVSASHLE 104  
RESULT 6  
US-08-110-786A-8  
Sequence 8, Application US/08110786A  
Patent No. 5443986  
GENERAL INFORMATION:  
APPLICANT: FAIRWEATHER, Neil Fraser  
APPLICANT: MAKOFF, Andrew Joseph  
TITLE OF INVENTION: Expression of tetanus toxin fragment C  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye P.C.  
STREET: 1100 No. 5443966th Glebe Road  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/110,786A  
FILING DATE: 23-AUG-1993 1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/777,337  
FILING DATE: 29-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB90/00943  
FILING DATE: 20-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8914122.0  
FILING DATE: 20 June 1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Mary J. Wilson  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 117-134  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-110-786A-8

Query Match 41.6%; Score 114; DB 1; Length 452;  
Best Local Similarity 95.5%; Pred. No. 2.4e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 16 LFNNFTVSFWLRPKVSASHLE 37  
Db 83 MFNNFTVSFWLRPKVSASHLE 104  
RESULT 7  
US-08-280-228-2

```

RESULT 2
US-08-488-351A-14
; Sequence 14, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Y1
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-14

Query Match 55.5%; Score 152; DB 2; Length 32;
Best Local Similarity 87.9%; Pred. No. 6.6e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 17 FNNFTVSFWLRVPKVSASHLEGLSP 49
Db 3 FNNFTVSFWLRVPKVSASHLE---HWSYGLRP 31

RESULT 3
PCT-US93-11703-64
; Sequence 64, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-11703-64

Query Match 41.6%; Score 114; DB 5; Length 31;
Best Local Similarity 95.5%; Pred. No. 8.8e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 6 MFNNFTVSFWLRVPKVSASHLE 27

RESULT 4
US-07-618-312A-2
; Sequence 2, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618,312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:46:20 ; Search time 14.8199 Seconds  
(without alignments)  
142.750 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHWSYGLRPGSSGSLFNNF.....VSASHLEGPLHWSYGLRPX 50

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap:\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pap:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	55.5	32	1	US-08-446-692-14
2	152	55.5	32	2	US-08-488-351A-14
3	114	41.6	31	5	PCT-US93-11703-64
4	114	41.6	452	1	US-07-618-312A-2
5	114	41.6	452	1	US-07-618-312A-4
6	114	41.6	452	1	US-08-110-786A-8
7	114	41.6	452	1	US-08-280-228-2
8	114	41.6	452	1	US-08-280-228-4
9	114	41.6	618	1	US-08-668-381A-5
10	114	41.6	853	4	US-08-913-880C-17
11	114	41.6	858	4	US-08-913-880C-16
12	114	41.6	860	4	US-08-913-880C-15
13	114	41.6	862	4	US-08-913-880C-14
14	114	41.6	865	4	US-08-913-880C-13
15	114	41.6	866	4	US-08-913-880C-12
16	114	41.6	874	4	US-08-913-880C-11
17	114	41.6	875	4	US-08-913-880C-10
18	114	41.6	1315	4	US-08-913-880C-1
19	112	40.9	21	1	US-07-610-525-1
20	112	40.9	21	2	US-08-661-052-12
21	112	40.9	21	2	US-08-460-502-8
22	112	40.9	21	2	US-08-724-774B-5
23	112	40.9	21	3	US-09-089-595-5
24	112	40.9	21	3	US-09-382-855-5
25	112	40.9	21	3	US-09-183-714B-5
26	112	40.9	21	3	US-09-188-082-12
27	112	40.9	21	3	US-09-171-969-10

28	112	40.9	21	4	US-09-364-088-12	Sequence 12, Appl
29	112	40.9	21	4	US-09-642-281-5	Sequence 5, Appl
30	112	40.9	21	4	US-09-102-716-12	Sequence 12, Appl
31	112	40.9	21	4	US-08-432-483A-3	Sequence 3, Appl
32	112	40.9	21	4	US-09-148-711A-8	Sequence 8, Appl
33	112	40.9	21	4	US-09-589-717-5	Sequence 5, Appl
34	112	40.9	21	4	US-08-945-289-3	Sequence 3, Appl
35	112	40.9	21	5	PCT-US93-11703-66	Sequence 66, Appl
36	107	39.1	22	1	US-08-446-692-5	Sequence 5, Appl
37	107	39.1	22	2	US-08-488-351A-5	Sequence 5, Appl
38	107	39.1	22	3	US-09-100-409A-41	Sequence 41, Appl
39	107	39.1	22	5	PCT-US95-13841-8	Sequence 8, Appl
40	102.5	37.4	699	3	US-08-694-865-16	Sequence 16, Appl
41	102.5	37.4	699	3	US-09-124-491-16	Sequence 16, Appl
42	102.5	37.4	699	4	US-09-383-912-16	Sequence 16, Appl
43	101	36.9	49	1	US-08-387-150-4	Sequence 4, Appl
44	101	36.9	49	2	US-08-694-865-4	Sequence 4, Appl
45	101	36.9	49	2	US-08-878-748-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-08-446-692-14  
; Sequence 14, Application US/08446692  
; Patent No. 5759551  
; GENERAL INFORMATION:  
; APPLICANT: Ladd, Anna  
; APPLICANT: Wang, Chang Y1  
; APPLICANT: Zamb, Timothy  
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maria C.H. Lin  
; CITY: New York  
; STATE: NY  
; COUNTRY: US  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,692  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria C.H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)415-8745  
; TELEFAX: (516)751-6849  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-446-692-14

Query Match 55.5%; Score 152; DB 1; Length 32;  
Best Local Similarity 87.9%; Pred. No. 6.6e-14;  
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 17 FNNFTVSWLRVPKVSASHLEGPLSHWSYGLRP 49  
|||||  
DB 3 FNNFTVSWLRVPKVSASHLE----HWSYGLRP 31

; CURRENT APPLICATION NUMBER: US/09/943,548  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 08/432,483  
; PRIOR FILING DATE: 1995-05-01  
; PRIOR APPLICATION NUMBER: PCT/US96/06147  
; PRIOR FILING DATE: 1996-05-01  
; PRIOR APPLICATION NUMBER: 08/945,289  
; PRIOR FILING DATE: 1997-10-17  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: helper T cell epitope of tetanus toxin  
US-09-943-548-3

Query Match 40.9%; Score 112; DB 9; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.7e-07;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 FNNFTVSFWLRVPKVSASHLE 37  
Db 1 FNNFTVSFWLRVPKVSASHLE 21

Search completed: September 24, 2003, 17:33:10  
Job time : 25.0693 secs

; PRIOR APPLICATION NUMBER: GB 9928530.6  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: GB 008658.7  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 907  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-130-973A-5

Query Match 41.6%; Score 114; DB 12; Length 907;  
Best Local Similarity 95.5%; Pred. No. 1.3e-05;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNFTVSVFVLRVPRKVSASHLE 37  
:|||||  
Db 536 MENFTVSVFVLRVPRKVSASHLE 557

RESULT 12  
US-10-130-973A-17  
; Sequence 17, Application US/10130973A  
; Publication No. US20030147895A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford  
; APPLICANT: Sutton, John  
; APPLICANT: Silman, Nigel  
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells  
; FILE REFERENCE: 1581.092000  
; CURRENT APPLICATION NUMBER: US/10/130,973A  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: PCT/GB00/04644  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: GB 9928530.6  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: GB 008658.7  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17  
; LENGTH: 1052  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-130-973A-17

Query Match 41.6%; Score 114; DB 12; Length 1052;  
Best Local Similarity 95.5%; Pred. No. 1.6e-05;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNFTVSVFVLRVPRKVSASHLE 37  
:|||||  
Db 681 MENFTVSVFVLRVPRKVSASHLE 702

RESULT 13  
US-10-130-973A-16  
; Sequence 16, Application US/10130973A  
; Publication No. US20030147895A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford  
; APPLICANT: Sutton, John  
; APPLICANT: Silman, Nigel  
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells  
; FILE REFERENCE: 1581.092000  
; CURRENT APPLICATION NUMBER: US/10/130,973A  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: PCT/GB00/04644

; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: GB 9928530.6  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: GB 008658.7  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 1112  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-130-973A-16

Query Match 41.6%; Score 114; DB 12; Length 1112;  
Best Local Similarity 95.5%; Pred. No. 1.7e-05;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNFTVSVFVLRVPRKVSASHLE 37  
:|||||  
Db 741 MENFTVSVFVLRVPRKVSASHLE 762

RESULT 14  
US-10-241-596-141  
; Sequence 141, Application US/10241596  
; Publication No. US20030166238A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbiological Research Authority  
; APPLICANT: The Speywood Laboratory Limited  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130003  
; CURRENT APPLICATION NUMBER: US/10/241,596  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: US 09/242,689  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: US 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB 9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB 9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 175  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 141  
; LENGTH: 1315  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-10-241-596-141

Query Match 41.6%; Score 114; DB 12; Length 1315;  
Best Local Similarity 95.5%; Pred. No. 2e-05;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNFTVSVFVLRVPRKVSASHLE 37  
:|||||  
Db 946 MENFTVSVFVLRVPRKVSASHLE 967

RESULT 15  
US-09-943-548-3  
; Sequence 3, Application US/09943548  
; Patent No. US20020042364A1  
; GENERAL INFORMATION:  
; APPLICANT: Rittershaus, Charles W.  
; APPLICANT: Thomas, Lawrence J.  
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY  
; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2

; PRIOR APPLICATION NUMBER: 60/055,615  
; PRIOR FILING DATE: 1997-08-14  
; PRIOR APPLICATION NUMBER: 60/065,236  
; PRIOR FILING DATE: 1997-11-13  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 463  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-09-816-467-2

Query Match 41.6%; Score 114; DB 11; Length 463;  
Best Local Similarity 95.5%; Pred. No. 6.3e-06;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37  
:|||||  
Db 94 MFNNFTVSFWLRVPKVSASHLE 115

## RESULT 8

US-10-130-973A-11  
; Sequence 11, Application US/10130973A  
; Publication No. US20030147895A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford  
; APPLICANT: Sutton, John  
; APPLICANT: Silman, Nigel  
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells  
; FILE REFERENCE: 1581.0920000  
; CURRENT APPLICATION NUMBER: US/10/130,973A  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: PCT/GB00/04644  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: GB 9928530.6  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: GB 008658.7  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 605  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-130-973A-11

Query Match 41.6%; Score 114; DB 12; Length 605;  
Best Local Similarity 95.5%; Pred. No. 8.5e-06;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37  
:|||||  
Db 234 MFNNFTVSFWLRVPKVSASHLE 255

## RESULT 9

US-10-130-973A-9  
; Sequence 9, Application US/10130973A  
; Publication No. US20030147895A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford  
; APPLICANT: Sutton, John  
; APPLICANT: Silman, Nigel  
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells  
; FILE REFERENCE: 1581.0920000  
; CURRENT APPLICATION NUMBER: US/10/130,973A  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: PCT/GB00/04644  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: GB 9928530.6

; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: GB 008658.7  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 665  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-130-973A-9

Query Match 41.6%; Score 114; DB 12; Length 665;  
Best Local Similarity 95.5%; Pred. No. 9.4e-06;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37  
:|||||  
Db 294 MFNNFTVSFWLRVPKVSASHLE 315

## RESULT 10

US-10-130-973A-3  
; Sequence 3, Application US/10130973A  
; Publication No. US20030147895A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford  
; APPLICANT: Sutton, John  
; APPLICANT: Silman, Nigel  
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells  
; FILE REFERENCE: 1581.0920000  
; CURRENT APPLICATION NUMBER: US/10/130,973A  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: PCT/GB00/04644  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: GB 9928530.6  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: GB 008658.7  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-130-973A-3

Query Match 41.6%; Score 114; DB 12; Length 882;  
Best Local Similarity 95.5%; Pred. No. 1.3e-05;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37  
:|||||  
Db 511 MFNNFTVSFWLRVPKVSASHLE 532

## RESULT 11

US-10-130-973A-5  
; Sequence 5, Application US/10130973A  
; Publication No. US20030147895A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford  
; APPLICANT: Sutton, John  
; APPLICANT: Silman, Nigel  
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells  
; FILE REFERENCE: 1581.0920000  
; CURRENT APPLICATION NUMBER: US/10/130,973A  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: PCT/GB00/04644  
; PRIOR FILING DATE: 2000-12-04

; LOCATION: (11)..(16)  
 ; OTHER INFORMATION: Spacer peptide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (17)..(31)  
 ; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor  
 ; OTHER INFORMATION: (Tentoxylisin)  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (32)..(37)  
 ; OTHER INFORMATION: Spacer peptide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (38)..(46)  
 ; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
 ; US-09-848-834A-19

Query Match 52.9%; Score 145; DB 9; Length 46;  
 Best Local Similarity 60.4%; Pred. No. 7.6e-11;  
 Matches 29; Conservative 6; Mismatches 9; Indels 4; Gaps 1;

Qy 2 HWSYGLRPGSSGSPSLFNFTVSFWLRVPKVSASHL-----GPSLHWSYGLRP 49  
 Db 2 HWSYGLRPGSSGSPSLQYIKANSKPGITELSS----GPSLHWSYGLRP 45

## RESULT 5

US-09-848-834A-17  
 ; Sequence 17, Application US/09848834A  
 ; Patent No. US20020076416A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aphton Corporation  
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens  
 ; FILE REFERENCE: 1102865-0047  
 ; CURRENT APPLICATION NUMBER: US/09/848,834A  
 ; CURRENT FILING DATE: 2001-05-04  
 ; PRIOR APPLICATION NUMBER: 60/202,328  
 ; PRIOR FILING DATE: 2000-05-05  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 17  
 ; LENGTH: 47  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn  
 ; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of  
 ; OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid seq  
 ; OTHER INFORMATION: uence 2-10 of the GnRH hormone  
 ; NAME/KEY: MOD\_RES  
 ; LOCATION: (1)..(1)  
 ; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
 ; NAME/KEY: MOD\_RES  
 ; LOCATION: (47)..(47)  
 ; OTHER INFORMATION: Amidated-glycine or glycineamide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(10)  
 ; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (11)..(18)  
 ; OTHER INFORMATION: Spacer peptide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (19)..(34)  
 ; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (35)..(38)  
 ; OTHER INFORMATION: Spacer peptide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (39)..(47)  
 ; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
 ; US-09-848-834A-17

Query Match 51.6%; Score 141.5; DB 9; Length 47;  
 Best Local Similarity 58.8%; Pred. No. 2.1e-10;  
 Matches 30; Conservative 1; Mismatches 11; Indels 9; Gaps 2;

Qy 2 HWSYGLRPGSSGSPSLFNFTVSFWLRVPKVSASHL-----EGPSLHWSYGLRP 49  
 Db 2 HWSYGLRPGSSGSPSL-----KLLSEIKGVIVHRLEGVEGSPSLHWSYGLRP 46

## RESULT 6

US-09-848-834A-20  
 ; Sequence 20, Application US/09848834A  
 ; Patent No. US20020076416A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aphton Corporation  
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens  
 ; FILE REFERENCE: 1102865-0047  
 ; CURRENT APPLICATION NUMBER: US/09/848,834A  
 ; CURRENT FILING DATE: 2001-05-04  
 ; PRIOR APPLICATION NUMBER: 60/202,328  
 ; PRIOR FILING DATE: 2000-05-05  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 20  
 ; LENGTH: 51  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of fur  
 ; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of PMS  
 ; OTHER INFORMATION: ium faiciparum circumsporoite (CSP) protein  
 ; NAME/KEY: MOD\_RES  
 ; LOCATION: (1)..(1)  
 ; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
 ; NAME/KEY: MOD\_RES  
 ; LOCATION: (51)..(51)  
 ; OTHER INFORMATION: Amidated glycine or glycineamide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(10)  
 ; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (11)..(16)  
 ; OTHER INFORMATION: Spacer peptide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (17)..(36)  
 ; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium faiciparum  
 ; OTHER INFORMATION: circumsporoite (CSP) protein  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (37)..(42)  
 ; OTHER INFORMATION: Spacer peptide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (43)..(51)  
 ; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
 ; US-09-848-834A-20

Query Match 50.5%; Score 138.5; DB 9; Length 51;  
 Best Local Similarity 54.7%; Pred. No. 5.4e-10;  
 Matches 29; Conservative 4; Mismatches 11; Indels 9; Gaps 2;

Qy 2 HWSYGLRPGSSGSPSLFNFTVSFWLRVPKVS-----SHLEGPSSLHWSYGLRP 49  
 Db 2 HWSYGLRPGSSGSPSLDEKKIA-----KMEKASSVFNWVNSSGSPSLHWSYGLRP 50

## RESULT 7

US-09-816-467-2  
 ; Sequence 2, Application US/09816467  
 ; Publication No. US20030004121A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: COEN, LAURENT  
 ; APPLICANT: PINZOLAS, ROSARIO OSTA  
 ; APPLICANT: BRULET, PHILIPPE  
 ; TITLE OF INVENTION: HYBRID PROTEINS THAT MIGRATE RETROGRADELY AND  
 ; TITLE OF INVENTION: TRANSNAPTICALLY INTO THE CNS  
 ; FILE REFERENCE: 03495-0174-01000  
 ; CURRENT APPLICATION NUMBER: US/09/816,467  
 ; CURRENT FILING DATE: 2001-03-26

OTHER INFORMATION: oxylysins  
NAME/KEY: PEPTIDE  
LOCATION: (38)..(41)  
OTHER INFORMATION: Spacer peptide  
NAME/KEY: PEPTIDE  
LOCATION: (42)..(50)  
OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-18

Query Match 99.3%; Score 272; DB 9; Length 50;  
Best Local Similarity 100.0%; Pred. No. 1.9e-26; Indels 0; Gaps 0;  
Matches 48; Conservative 0; Mismatches 0;

Qy 2 HWSYGLRPGSGPSLFNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 49  
|||||  
Db 2 HWSYGLRPGSGPSLFNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 49

## RESULT 2

US-09-848-834A-14  
Sequence 14, Application US/09848834A  
Patent No. US20020076416A1

GENERAL INFORMATION:  
APPLICANT: Aphton Corporation  
TITLE OF INVENTION: Chimeric Peptide Immunogens  
FILE REFERENCE: 1102865-0047  
CURRENT APPLICATION NUMBER: US/09/848,834A  
CURRENT FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: 60/202,328  
PRIOR FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.0

SEQ ID NO 14  
LENGTH: 37  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the human GnRH hormone  
OTHER INFORMATION: GnRH hormone linked by a spacer to amino acid sequence 947-967 of the Tetanus toxoid precursor (Tentoxylisin)

NAME/KEY: MOD\_RES  
LOCATION: (1)..(1)  
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
NAME/KEY: PEPTIDE

LOCATION: (1)..(10)  
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
NAME/KEY: PEPTIDE

LOCATION: (11)..(16)  
OTHER INFORMATION: Spacer peptide  
NAME/KEY: PEPTIDE

LOCATION: (17)..(37)  
OTHER INFORMATION: (Tentoxylisin)

US-09-848-834A-14  
Query Match 72.6%; Score 199; DB 9; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.3e-17;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSGPSLFNNFTVFWLRVPKVSASHLE 37  
|||||  
Db 2 HWSYGLRPGSGPSLFNNFTVFWLRVPKVSASHLE 37

## RESULT 3

US-09-848-834A-10  
Sequence 10, Application US/09848834A  
Patent No. US20020076416A1

GENERAL INFORMATION:  
APPLICANT: Aphton Corporation  
TITLE OF INVENTION: Chimeric Peptide Immunogens  
FILE REFERENCE: 1102865-0047  
CURRENT APPLICATION NUMBER: US/09/848,834A

CURRENT FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: 60/202,328  
PRIOR FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 10  
LENGTH: 34  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of the Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to amino acid sequence 2-10 of the GnRH hormone  
NAME/KEY: MOD\_RES  
LOCATION: (1)..(1)

OTHER INFORMATION: Amidated phenylalanine  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(21)

OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor  
OTHER INFORMATION: (Tentoxylisin)

NAME/KEY: PEPTIDE  
LOCATION: (22)..(25)

OTHER INFORMATION: Spacer peptide  
NAME/KEY: PEPTIDE  
LOCATION: (26)..(34)

OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone  
NAME/KEY: MOD\_RES  
LOCATION: (34)..(34)

OTHER INFORMATION: Amidated glycine or glycylamide  
US-09-848-834A-10

Query Match 67.5%; Score 185; DB 9; Length 34;  
Best Local Similarity 100.0%; Pred. No. 6.4e-16;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 49  
|||||  
Db 1 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 33

## RESULT 4

US-09-848-834A-19  
Sequence 19, Application US/09848834A  
Patent No. US20020076416A1

GENERAL INFORMATION:  
APPLICANT: Aphton Corporation  
TITLE OF INVENTION: Chimeric Peptide Immunogens  
FILE REFERENCE: 1102865-0047

CURRENT APPLICATION NUMBER: US/09/848,834A  
CURRENT FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: 60/202,328

PRIOR FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.0

SEQ ID NO 19  
LENGTH: 46  
TYPE: PRT

ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the human GnRH hormone linked by a spacer to amino acid sequence 830-844 of Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to amino acid sequence 1-10 of GnRH

NAME/KEY: MOD\_RES  
LOCATION: (1)..(1)

OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
NAME/KEY: MOD\_RES  
LOCATION: (46)..(46)

OTHER INFORMATION: Amidated glycine or glycylamide  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(10)

OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
NAME/KEY: PEPTIDE

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:57:10 ; Search time 25.0693 Seconds  
(without alignments)  
301.778 Million cell updates/sec

Title: US-09-848-834A-18  
Perfect score: 274  
Sequence: 1 XHWSYGLRPGSGPSLFNFF.....VSASHLEGPSLHWSVGLRPX 50

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues  
Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	272	99.3	50	9	US-09-848-834A-18
2	199	72.6	37	9	US-09-848-834A-14
3	185	67.5	34	9	US-09-848-834A-10
4	145	52.9	46	9	US-09-848-834A-19
5	141.5	51.6	47	9	US-09-848-834A-17
6	138.5	50.5	51	9	US-09-848-834A-20
7	114	41.6	463	11	US-09-816-467-2
8	114	41.6	605	12	US-10-130-973A-11
9	114	41.6	665	12	US-10-130-973A-9
10	114	41.6	882	12	US-10-130-973A-3
11	114	41.6	907	12	US-10-130-973A-5
12	114	41.6	1052	12	US-10-130-973A-17
13	114	41.6	1112	12	US-10-130-973A-16
14	114	41.6	1315	12	US-10-241-596-141
15	112	40.9	21	9	US-09-943-548-3

Sequence 4, Appli	21	9	US-09-848-834A-4
Sequence 6, Appli	21	10	US-09-785-215-6
Sequence 2, Appli	21	12	US-09-405-986-2
Sequence 6, Appli	21	12	US-10-223-809A-6
Sequence 5, Appli	21	12	US-10-261-208-5
Sequence 6, Appli	21	15	US-10-204-362-6
Sequence 3, Appli	21	15	US-10-339-522-3
Sequence 8, Appli	21	15	US-10-223-711-8
Sequence 4, Appli	22	11	US-09-865-294-4
Sequence 13, Appli	695	11	US-09-305-924-13
Sequence 4, Appli	9	9	US-09-019-010-4
Sequence 11, Appl	49	11	US-09-305-924-11
Sequence 35, Appl	49	11	US-09-964-201A-35
Sequence 34, Appl	40	11	US-09-964-201A-34
Sequence 20, Appl	41	11	US-09-964-201A-34
Sequence 12, Appl	20	14	US-10-044-034-20
Sequence 15, Appl	22	11	US-09-747-802-12
Sequence 13, Appl	31	9	US-09-848-834A-15
Sequence 16, Appl	33	9	US-09-848-834A-16
Sequence 26, Appl	36	9	US-09-848-834A-16
Sequence 29, Appl	35	85	31.0
Sequence 30, Appl	20	11	US-09-964-201A-29
Sequence 31, Appl	20	11	US-09-964-201A-30
Sequence 43, Appl	20	11	US-09-964-201A-31
Sequence 44, Appl	14	10	US-09-909-460-43
Sequence 9, Appli	15	10	US-09-909-460-44
Sequence 11, Appl	31	9	US-09-848-834A-9
Sequence 12, Appl	28	9	US-09-848-834A-11
Sequence 4, Appli	73	26.6	33
Sequence 18, Appl	33	9	US-09-848-834A-12
	71	25.9	17
	44		11
	24.5		449
			11
			US-09-910-186A-18

ALIGNMENTS

RESULT 1  
US-09-848-834A-18  
; Sequence 18, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 18  
; LENGTH: 50  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of hr  
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the  
; OTHER INFORMATION: anus toxoid precursor (tetoxylisin) protein linked by a spa  
; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: MOD\_RES  
; LOCATION: (50)..(50)  
; OTHER INFORMATION: Amidated glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(37)  
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor

```
DR InterPro; IPR006025; Zn_MTpeptdse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR ProDom; PD001963; Bontotoxylisin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0
FT CHAIN 1 421
FT METAL 422 1250
FT METAL 211 211
FT ACT_SITE 212 212
FT METAL 215 215
FT DISULFID 411 425
FT CONFLICT 176 176
FT CONFLICT 197 197
FT CONFLICT 339 339
FT CONFLICT 772 772
FT CONFLICT 962 963
FT CONFLICT 966 966
FT CONFLICT 1194 1194
SQ SEQUENCE 1250 AA; 143712 MW; D9FCE26DDA041EB4 CRC64;

Query Match 21.48; Score 58.5; DB 1; Length 1250;
Best Local Similarity 23.68; Pred. No. 26;
Matches 13; Conservative 9; Mismatches 4; Indels 29; Gaps 2;

Qy 17 FNNFTVSFWLRVP-----KVSASHLEGPSLHWY 45
Db 911 YKNEISISFWIRIPYNDKIVNNNEYTIINCMDNNSGWKVSLNHNE---IITWF 962

RESULT 15
ID GONL_SHEEP STANDARD; PRT; 61 AA.
AC Q28588;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
DE GNRH1 OR GNRH OR LHRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE OF 1-10.
RX MEDLINE=72094314; PubMed=4550508;
RA Burgess R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
RA Fellows R., Blackwell R., Vale W., Guillemin R.;
RT "Primary structure of the ovine hypothalamic luteinizing hormone-
RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
RT spectrometry-decapeptide-Edman degradation).";
RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GNRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U02517; AAA03433.1; -.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadolibereinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pyrrolidone carboxylic acid.
FT NON_TER 1
FT CHAIN 1 >61
FT PEPTIDE 1 10
FT PEPTIDE 14 >61
FT ACT_SITE 3 3
FT MOD_RES 1 1
FT MOD_RES 10 10
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;

Query Match 21.28; Score 58; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
Db 2 HWSYGLRPG 10

Search completed: September 24, 2003, 17:41:21
Job time : 9.37978 secs
```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; X01059; CAA25526.1; -
CC EMBL; M12578; AAA35916.1; -
CC EMBL; X15215; CAA33285.1; -
CC PIR; S05308; RHUG.
DR Genew; HGNC:4419; GNRH1.
DR MIM; 152760; -
DR GO; GO:0003625; C:soluble fraction; TAS.
DR GO; GO:0005183; F:lutalizing hormone-releasing factor activity; TAS.
DR GO; GO:0007267; P:cell-cell signaling; TAS.
DR GO; GO:0007275; P:development; TAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBRNI.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmaceutical; Signal; Polymorphism;
KW Pyrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT VARIANT 16 16
FT VARIANT /FTID=VAR_013943.
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match 21.5%; Score 59; DB 1; Length 92;
Best Local Similarity 52.4%; Pred. No. 1.3;
Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 2 HWSYGLRPGS--SGPSLFNNF 20
DB 25 HWSYGLRPGGRDAENLIDSF 45
||||||| :|:|
||||||| :|:|

RESULT 14
BXE_CLOBO STANDARD; PRT; 1250 AA.
AC Q00496;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
DE (Bontoxilysin E).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beluga;
RX MEDLINE=92181428; PubMed=1543481;
RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT "Sequences of the botulinum neurotoxin E derived from Clostridium
RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT ATCC 43181 and ATCC 43755).";
RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92174922; PubMed=1541280;
RA Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
RA "The complete amino acid sequence of the Clostridium botulinum type-E
RT neurotoxin, derived by nucleotide-sequence analysis of the encoding

```

```

RT gene.";
RL Eur. J. Biochem. 204:657-667(1992).
RN [3]
RP SEQUENCE OF 1-251 FROM N.A.
RX MEDLINE=90264400; PubMed=2160960;
RA Binz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;
RT "The complete sequence of botulinum neurotoxin type A and comparison
RT with other clostridial neurotoxins.";
RL J. Biol. Chem. 265:9153-9158(1990).
RN [4]
RP SEQUENCE OF 1-13.
RX MEDLINE=85197963; PubMed=3888113;
RA Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequences of botulinum neurotoxins types B and
RT E.";
RL Arch. Biochem. Biophys. 238:544-548(1985).
RN [5]
RP SEQUENCE OF 419-426.
RX MEDLINE=90344918; PubMed=2116911;
RA Gimenez J.A., Dasgupta B.R.;
RT "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
RT reveals the site trypsin nicks and homology with tetanus
RT neurotoxin.";
RL Biochimie 72:213-217(1990).
RN [6]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94063091; PubMed=8243676;
RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA Benfenati F., Wilson M.C., Montecucco C.;
RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RT COOH-terminal peptide bonds.";
RL FEBS Lett. 335:99-103(1993).
RN [7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94124495; PubMed=8294407;
RA Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA Jahn R., Niemann H.;
RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
RL J. Biol. Chem. 269:1617-1620(1994).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-|-ILE-
CC 181 BOND IN SNAP-25.
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; X62089; CAA43999.1; -
CC EMBL; X62863; CAA44558.1; -
CC PIR; S21178; S21178.
CC HSSP; P10845; 3BTA.
CC MEROPS; M27.002; -
CC InterPro; IPR000395; Bontoxilysin.

```

```

GONI_MACMU          STANDARD;          PRT;          67 AA.
ID  GONI_MACMU          STANDARD;          PRT;          92 AA.
AC  P55247;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Progonadolibin I precursor [Contains: Gonadolibin I (LH-RH I)
DE  (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE  hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE  (Fragment).
DE  GNRH1 OR GNRH OR LHRH.
GN  Macaca mulatta (Rhesus macaque).
OS  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC  Cercopitheciinae; Macaca.
OX  NCBI_TaxID=9544;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Hypothalamus;
RX  MEDLINE=95124501; PubMed=7545971;
RA  Ma Y.J., Costa M.E., Ojeda S.R.;
RT  "Developmental expression of the genes encoding transforming growth
RT  factor alpha and its receptor in the hypothalamus of female rhesus
RT  macaques.";
RL  Neuroendocrinology 60:346-359(1994).
CC  -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC  THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC  HORMONES.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- SIMILARITY: Belongs to the GNRH family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; S75918; AAB33096.1; -.
DR  PIR; I78541.
DR  InterPro; IPR002012; GNRH.
DR  InterPro; IPR004079; GonadolibinI.
DR  Pfam; PF00446; Gnrh; 1.
DR  PRINTS; PS01541; GONADOLIBRN1.
DR  PROSITE; PS00473; GNRH; 1.
KW  Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW  Signal; Pyrrolidone carboxylic acid.
FT  NO_TER  1 1
FT  SIGNAL  <1 5 BY SIMILARITY.
FT  CHAIN  6 >67 PROGONADOLIBIN I.
FT  PEPTIDE 6 15 GONADOLIBIN I.
FT  PEPTIDE 19 >67 GNRH-ASSOCIATED PEPTIDE I.
FT  ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT  ACT_SITE 8 8 ACTIVITY (BY SIMILARITY).
FT  MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
FT  MOD_RES 15 15 AMIDATION).
FT  MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
FT  NON_TER 67 67 SIMILARITY).
SQ  SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 21.5%; Score 59; DB 1; Length 67;
Best Local Similarity 52.4%; Pred. No. 0.92;
Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

Qy 2 HWSYGLRPGS--SQPSLFNNF 20
Db 7 HWSYGLRPGGKRDAENLMSDF 27
      |||||
      :|:|
      :|:|

RESULT 13
GONI_HUMAN

```

```

ID  GONI_HUMAN          STANDARD;          PRT;          92 AA.
AC  P01148;
DT  21-JUL-1986 (Rel. 01, Created)
DT  01-APR-1988 (Rel. 07, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Progonadolibin I precursor [Contains: Gonadolibin I (LH-RH I)
DE  (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE  hormone I) (GnRH I) (Luliberin I) (Gonadorelin I); GnRH-associated
DE  peptide I].
DE  GNRH1 OR GNRH OR LHRH.
GN  Homo sapiens (Human).
OS  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  MEDLINE=89366682; PubMed=2671939;
RX  Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT  "The complete nucleotide sequence of the human gonadotropin-releasing
RT  hormone gene.";
RL  Nucleic Acids Res. 17:6403-6403(1989).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=86094338; PubMed=2867548;
RA  Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT  "Isolation of the gene and hypothalamic cDNA for the common precursor
RT  of gonadotropin-releasing hormone and prolactin release-inhibiting
RT  factor in human and rat.";
RL  Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN  [3]
RP  SEQUENCE FROM N.A., AND VARIANT SER-16.
RX  MEDLINE=85012739; PubMed=6090951;
RA  Seeburg P.H., Adelman J.P.;
RT  "Characterization of cDNA for precursor of human luteinizing hormone
RT  releasing hormone.";
RL  Nature 311:666-668(1984).
RN  [4]
RP  SEQUENCE OF 24-33.
RX  MEDLINE=83126573; PubMed=6760865;
RA  Tan L., Rousseau P.;
RT  "The chemical identity of the immunoreactive LHRH-like peptide
RT  biosynthesized in the human placenta.";
RL  Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
RN  [5]
RP  VARIANT SER-16.
RX  MEDLINE=99318093; PubMed=10391209;
RA  Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA  Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
RA  Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA  Lander E.S.;
RT  "Characterization of single-nucleotide polymorphisms in coding regions
RT  of human genes.";
RL  Nat. Genet. 22:231-238(1999).
RN  [6]
RP  ERRATUM.
RA  Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA  Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
RA  Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA  Lander E.S.;
RL  Nat. Genet. 23:373-373(1999).
CC  -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC  THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC  HORMONES.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC  Lutrepulse or Lutrelef (Ferring Pharmaceuticals) and Relisorm
CC  (Serono).
CC  -1- SIMILARITY: Belongs to the GNRH family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way

```

```

RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.;"
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [2]
RN SEQUENCE FROM N.A.
RP
RX MEDLINE=89384661; PubMed=2476669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and
RT hypothalamic expression.;"
RL Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RN SEQUENCE FROM N.A.
RP
RC TISSUE=Thymus;
RX MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA.;"
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RN SEQUENCE OF 1-47 FROM N.A.
RP
RC TISSUE=Heart;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus.;"
RL Science 235:1514-1517(1987).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S50870; AAB24572.1; -.
DR EMBL; M12579; AAA41263.1; -.
DR EMBL; M31670; AAA41264.1; -.
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -.
DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
DR PIR; A40147; RHRTG.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT ACTIVITY
FT MOD_RES 24 24
FT MOD_RES 33 33
FT AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;
Query Match 22.4%; Score 61.5; DB 1; Length 92;
Best Local Similarity 51.7%; Pred. No. 0.64;
Matches 15; Conservative 3; Mismatches 4; Indels 7; Gaps 2;
QY 28 VPKVSAS-----HLEG-PSLHWSYGLRP 49

```

```

Db 4 IPKLMAAVVLLTVLCGCSQHSYGLRP 32
      ||| | : ||| | ||| |||||
      4 IPKLMAAVVLLTVLCGCSQHSYGLRP 32
RESULT 11
GONI_TUPGB STANDARD; PRT; 92 AA.
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
GN GnRH1 OR GnRH.
OS Tupaia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxID=9396;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kastan T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGnRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GnRH gene expression in a
RT placental mammal.;"
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U63326; AAB16837.1; -.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT ACTIVITY
FT MOD_RES 24 24
FT MOD_RES 33 33
FT AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
FT SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;
Query Match 21.9%; Score 60; DB 1; Length 92;
Best Local Similarity 52.4%; Pred. No. 0.97;
Matches 11; Conservative 4; Mismatches 4; Indels 2; Gaps 1;
QY 2 HWSYGLRPGS--SGPSLFNNF 20
      ||||| : :| :|
      25 HWSYGLRPGGKRNENLIDSF 45
RESULT 12

```



```
CC METAL      233    233 ZINC (CATALYTIC) (BY SIMILARITY).
CC DISULFID   436    445 INTERCHAIN (PROBABLE).
FT CONFLICT   29     29 T -> M (IN REF. 4).
CC CONFLICT   217    217 R -> G (IN REF. 2).
FT CONFLICT   224    224 A -> S (IN REF. 2).
FT CONFLICT   463    463 S -> R (IN REF. 4).
SQ SEQUENCE 1290 AA; D21746E2C024DF43 CRC64;

Query Match          23.0%; Score 63; DB 1; Length 1290;
Best Local Similarity 62.5%; Pred. No. 7.4;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 15 SLFNNTVSWLRVPK 30
Db 920 SVFLDPSFWIRPK 935

RESULT 7
BXCL_CLOBO ID BXCL_CLOBO STANDARD; PRT; 1290 AA.
AC PI8640;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BoNT/C1)
DE (Bontoxilysin Cl.)
DE Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxId=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90370487; PubMed=2204031;
RA Hauser D., Eklund M.W., Kurazono H., Binz T., Niemann H., Gill D.M., Boquet P., Popoff M.R.;
RL "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";
RT Nucleic Acids Res. 18:4924-4924(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Type C Stockholm / C-ST;
RX MEDLINE=91024998; PubMed=2222445;
RA Kimura K., Fujii N., Tsuzuki K., Murakami T., Indo H., Yokosawa N., Takeshi K., Syuto B., Oguma K.;
RT "The complete nucleotide sequence of the gene coding for botulinum type C1 toxin in the C-ST phage genome.";
RL Biochem. Biophys. Res. Commun. 171:1304-1311(1990).
RN [3]
RP SEQUENCE OF 2-25.
RC STRAIN-Type C Stockholm / C-ST;
RX MEDLINE=88153072; PubMed=2450068;
RA Tsuzuki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K., Oguma K.;
RT "Establishment of a monoclonal antibody recognizing an antigenic site common to Clostridium botulinum type B, C1, D, and E toxins and tetanus toxin.";
RL Infect. Immun. 56:898-902(1988).
RN [4]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94038966; PubMed=7901002;
RA Blas J., Chapman E.K., Yamasaki S., Binz T., Niemann H., Jahn R.;
RT "Botulinum neurotoxin C1 blocks neurotransmitter release by means of cleaving HPC-1/syntaxin.";
RL EMBO J. 12:4821-4828(1993).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC ENDOPEPTIDASE THAT CLEAVES SYNTAXIN.
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No detected action on small molecule substrates.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
```

DR MEROPS; M27.002; -.  
 DR InterPro; IPR000395; Bontoxilysin.  
 DR Pfam; PF01742; Peptidase\_M27; 1.  
 DR ProDom; PD001963; Bontoxilysin; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.  
 FT CHAIN 1 436  
 FT METAL 437 1274 BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.  
 FT ACT\_SITE 227 227 BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.  
 FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 231 231 BY SIMILARITY.  
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT DISULFID 429 445 INTERCHAIN (PROBABLE).  
 SQ SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;

Query Match 23.0%; Score 63; DB 1; Length 1274;  
 Best Local Similarity 56.2%; Pred. No. 7.3;  
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 15 SLFNFTVSWLRVPK 30  
 I : ||:||||:|  
 Db 928 SRYNFSISFWRIKP 943

RESULT 6  
 BXB\_CLOBO STANDARD; PRT; 1290 AA.  
 AC P10844; P10843;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BONT/B)  
 DE (Bontoxilysin B).  
 GN BOTB.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1491;  
 RN [1]  
 RX SEQUENCE FROM N.A.  
 RX MEDLINE=92384550; PubMed=1514783;  
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,  
 RA Minton N.P.;  
 RT "Molecular cloning of the Clostridium botulinum structural gene  
 RT encoding the type B neurotoxin and determination of its entire  
 RT nucleotide sequence.";  
 RL Appl. Environ. Microbiol. 58:2345-2354(1992).  
 RN [2]  
 RP SEQUENCE OF 35-245 FROM N.A.  
 RP STRAIN-NCTC 7273;  
 RA Szabo E.A., Pemberton J.M., Desmarchelier P.M.;  
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 633-993 FROM N.A.  
 RP STRAIN-NCTC 7273;  
 RX MEDLINE=94013372; PubMed=8408542;  
 RA Campbell K., East A.K., Collins M.D.;  
 RT "Gene probes for identification of the botulin neurotoxin gene and  
 RT specific identification of neurotoxin types B, E, and F.";  
 RL J. Clin. Microbiol. 31:2255-2262(1993).  
 RN [4]  
 RP SEQUENCE OF 1-44 AND 441-466.  
 RP STRAIN=657;  
 RC MEDLINE=89000987; PubMed=3139097;  
 RX Dasgupta B.R., Datta A.;  
 RA "Botulinum neurotoxin type B (strain 657): partial sequence and  
 RT similarity with tetanus toxin.";  
 RL Biochimie 70:811-817(1988).  
 RN [5]  
 RP SEQUENCE OF 1-16 AND 441-458.  
 RP STRAIN=OKRA;  
 RC MEDLINE=85197963; PubMed=3888113;  
 RX Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;

RT "Partial amino acid sequences of botulinum neurotoxins types B and  
 E.";  
 RL Arch. Biochem. Biophys. 238:544-548(1985).  
 RN [6]  
 RP IDENTIFICATION AS ZINC-PROTEASE.  
 RX MEDLINE=93054694; PubMed=1429690;  
 RA Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;  
 RT "Botulinum neurotoxins are zinc proteins.";  
 RL J. Biol. Chem. 267:23479-23483(1992).  
 RN [7]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=93063293; PubMed=1331807;  
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,  
 RA Dasgupta B.R., Montecucco C.;  
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release  
 RT by proteolytic cleavage of synaptobrevin.";  
 RL Nature 359:832-835(1992).  
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF  
 CC SYNAPTOSOMAL-2.  
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No  
 CC detected action on small molecule substrates.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A  
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,  
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL  
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF  
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M81186; AAA23211.1; -;  
 CC EMBL; Z11934; CAA77991.1; -;  
 CC EMBL; X70817; CAA50148.1; -;  
 DR PIR; A48940; A48940.  
 DR PDB; 1EPW; 01-NOV-00.  
 DR PDB; 1F31; 01-NOV-00.  
 DR PDB; 1F82; 16-AUG-00.  
 DR PDB; 1F83; 16-AUG-00.  
 DR PDB; 1FOH; 06-DEC-00.  
 DR PDB; 1G9A; 13-NOV-02.  
 DR PDB; 1G9B; 13-NOV-02.  
 DR PDB; 1G9C; 13-NOV-02.  
 DR PDB; 1G9D; 13-NOV-02.  
 DR PDB; 1I1E; 21-NOV-01.  
 DR MEROPS; M27.002; -;  
 DR InterPro; IPR000395; Bontoxilysin.  
 DR InterPro; IPR006025; Zn\_MTPeptidse.  
 DR Pfam; PF01742; Peptidase\_M27; 1.  
 DR PRINTS; PR00760; BONTOXILYSIN.  
 DR ProDom; PD001963; Bontoxilysin; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;  
 KW 3D-structure.  
 FT INIT\_MET 0  
 FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.  
 FT CHAIN 441 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.  
 FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 230 230 BY SIMILARITY.



RL FNBS Lett. 376:41-44(1995).  
 RP [6] SEQUENCE OF 1-16.  
 RX MEDLINE=84178501; PubMed=6370252;  
 RA Schmidt J.J., Sartyworthy V., Dasgupta B.R.;  
 RL "Partial amino acid sequence of the heavy and light chains of  
 RL botulinum neurotoxin type A.";  
 RL Biochem. Biophys. Res. Commun. 119:900-904(1984).  
 RP [7]  
 RP SEQUENCE OF 1-46.  
 RA Dasgupta B.R., Foley J., Niece R.;  
 RL "Partial sequence of the light chain of botulinum neurotoxin type A.";  
 RL Biochemistry 26:4182-4182(1987).  
 RP [8]  
 RP SEQUENCE OF 1-5 AND 44-456.  
 RX MEDLINE=91120847; PubMed=2126206;  
 RA Dasgupta B.R., Dekleva M.L.;  
 RL "Botulinum neurotoxin type A: sequence of amino acids at the  
 RL N-terminus and around the nicking site.";  
 RL Biochimie 72:661-664(1990).  
 RP [9]  
 RP SEQUENCE OF 448-464 AND 872-895.  
 RX MEDLINE=89024662; PubMed=3178218;  
 RA Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;  
 RL "Botulinum neurotoxin type A: cleavage of the heavy chain into two  
 RL halves and their partial sequences.";  
 RL Arch. Biochem. Biophys. 266:142-151(1988).  
 RP [10]  
 RP SEQUENCE OF 448-482.  
 RX MEDLINE=85285016; PubMed=3896784;  
 RA Shone C.C., Hambleton P., Melling J.;  
 RL "Inactivation of Clostridium botulinum type A neurotoxin by trypsin  
 RL and purification of two tryptic fragments. Proteolytic action near  
 RL the COOH-terminus of the heavy subunit destroys toxin-binding  
 RL activity.";  
 RL Eur. J. Biochem. 151:75-82(1985).  
 RP [11]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=94063091; PubMed=8243676;  
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,  
 RA Benfenati F., Wilson M.C., Montecucco C.;  
 RL "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct  
 RL COOH-terminal peptide bonds.";  
 RL FNBS Lett. 335:99-103(1993).  
 RP [12]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=94124495; PubMed=8294407;  
 RA Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,  
 RA Jahn R., Niemann H.;  
 RL "Proteolysis of SNAP-25 by types E and A botulin neurotoxins.";  
 RL J. Biol. Chem. 269:1617-1620(1994).  
 RP [13]  
 RP MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.  
 RX MEDLINE=21556941; PubMed=11700044;  
 RA Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.;  
 RL "Site-directed mutagenesis identifies active-site residues of the  
 RL light chain of botulinum neurotoxin type A.";  
 RL Biochem. Biophys. Res. Commun. 288:1231-1237(2001).  
 RP [14]  
 RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).  
 RX MEDLINE=98455071; PubMed=9787350;  
 RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;  
 RL "Crystal structure of botulinum neurotoxin type A and implications  
 RL for toxicity.";  
 RL Nat. Struct. Biol. 5:898-902(1998).  
 CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin  
 CC binds with high affinity to peripheral neuronal presynaptic  
 CC membrane, is then internalized by receptor-mediated endocytosis.  
 CC The C-terminus of the heavy chain (H) is responsible for the  
 CC adherence of the toxin to the cell surface while the N-terminus  
 CC mediates transport of the light chain from the endocytic vesicle  
 CC to the cytosol. After translocation, the light chain (L)  
 CC hydrolyzes the 197-Gln-1 Arg-198 bond in SNAP-25, thereby blocking

neurotransmitter release. Inhibition of acetylcholine release  
 results in flaccid paralysis, with frequent heart or respiratory  
 failure.  
 -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
 neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No  
 detected action on small molecule substrates.  
 -1- COFACTOR: Binds 1 zinc ion per subunit.  
 -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a  
 heavy chain (H).  
 -1- SUBCELLULAR LOCATION: Secreted.  
 -1- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for  
 the treatment of strabismus and blepharospasm associated with  
 dystonia and cervical dystonia. Also used for the treatment of  
 hemifacial spasm and a number of other neurological disorders  
 characterized by abnormal muscle contraction.  
 -1- MISCELLANEOUS: There are seven antigenically distinct forms of  
 botulinum neurotoxin: Types A, B, C1, D, E, F, and G.  
 -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 -1- DATABASE: NAME-BOTOX product information Web site:  
 WWW="http://www.botox.com/index.jsp?hp=productinfo".  
 -1- DATABASE: NAME-Protein Spotlight;  
 NOTE-Issue 19 of February 2002;  
 WWW="http://www.expasy.org/spotlight/articles/spl019.html".  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; X52066; CAA36289.1; -;  
 DR EMBL; M30196; AAA23262.1; -;  
 DR EMBL; X92973; CAA63551.1; -;  
 DR EMBL; D67030; BAA11051.1; -;  
 DR EMBL; M27892; AAA23269.1; -;  
 DR PIR; A35294; BTCLAB.  
 DR PDB; 3BTA; 01-OCT-99.  
 DR MEROPS; M27.002; -;  
 DR InterPro; IPR000395; Bontoxilysin.  
 DR InterPro; IPR006025; Zn\_MTPeptide.  
 DR Pfam; PF01742; Peptidase\_M27; 1.  
 DR ProDom; PD001963; Bontoxilysin; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;  
 KW Pharmaceutical; 3D-structure.  
 FT INIT\_MET 0  
 FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.  
 FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.  
 FT METAL 222 222 ZINC (CATALYTIC).  
 FT ACT\_SITE 223 223  
 FT METAL 226 226 ZINC (CATALYTIC).  
 FT METAL 261 261 ZINC (CATALYTIC).  
 FT DISULFID 429 453 INTERCHAIN.  
 FT DISULFID 1234 1279  
 FT TRANSMEM 626 646  
 FT TRANSMEM 655 675 POTENTIAL.  
 FT VARIANT 26 26 V -> A.  
 FT VARIANT 261 261 E -> A; DRASTIC DECREASE IN ENZYMIC  
 FT MUTAGEN 261 261 ACTIVITY.  
 FT MUTAGEN 265 265 F -> A; DECREASE IN ENZYMIC ACTIVITY.  
 FT MUTAGEN 365 365 Y -> A; DECREASE IN ENZYMIC ACTIVITY.  
 FT CONFLICT 1 1 P -> Q (IN REF. 1).  
 FT CONFLICT 479 479 E -> P (IN REF. 9).  
 FT CONFLICT 875 875 T -> L (IN REF. 8).  
 FT CONFLICT 891 891 S -> K (IN REF. 8).  
 SQ SEQUENCE 1295 AA; 149322 MW; 858342f754862579 CRC64;  
 Query Match 23.7%; Score 65; DB 1; Length 1295;  
 Best Local Similarity 56.2%; Pred. No. 4.2;  
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

FT METAL          229      229      ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE       230      230      BY SIMILARITY.
FT FT METAL       233      233      ZINC (CATALYTIC) (BY SIMILARITY).
FT DT SULFID      435      449      INTERCHAIN (PROBABLE).
SQ SEQUENCE       1296 AA; 149013 MW; DC8E47E15F665C31 CRC64;

Query Match                24.5%; Score 67; DB 1; Length 1296;
Best Local Similarity 39.1%; Pred. No. 2.4;
Matches    9; Conservative 10; Mismatches   4; Indels   0; Gaps   0;

QY      15 SLENNETVSPWLVRPKVSASHLE 37
         |::|::|::|::|::|::|::|::|
DB      927 SMEDNFSINFWRTPKYNNDIQ 949

RESULT 3
BXAL_CLOBO
ID     BXAL_CLOBO STANDARD; PRT; 1295 AA.
AC AC10845; P01561; P18639;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE (Sontoxilysin A) [BOROX] [Contains: Botulinum neurotoxin A, light-chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxId=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 2916;
RX MEDLINE=90235864; PubMed=2185020;
RA Thompson D.E., Brehm J.K., Oultam J.D., Swinfield T.-J.,
R Shone C.C., Atkinson T., Melling J., Minton N.P.;
RT "The complete amino acid sequence of the Clostridium botulinum type A
RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
RT gene.";
RL Eur. J. Biochem. 189:73-81(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GZA;
RX MEDLINE=90264400; PubMed=2160960;
RA Binz B., Kuwarzo H., Wille M., Frevent J., Wernars K., Niemann H.;
RT "The complete sequence of botulinum neurotoxin type A and comparison
RT with other clostridial neurotoxins.";
RL J. Biol. Chem. 265:9153-9158(1990).
RN [3]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=GZA;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium
RT botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
RN [4]
RP SEQUENCE OF 1-34 FROM N.A.
RC STRAIN=Hall;
RX MEDLINE=89350959; PubMed=2669749;
RA Betley M.J., Somers E., Dasgupta B.R.;
RT "Characterization of botulinum type A neurotoxin gene: delineation of
RT the N-terminal encoding region.";
RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
RN [5]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=Type A NIH;
RX MEDLINE=96096783; PubMed=8521962;
RA Fujita R., Fujiwara Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin
RT components of Clostridium botulinum type A progenitor toxins.";
```

RL Eur. J. Biochem. 188:39-45(1990).  
RN [6]  
RP PARTIAL SEQUENCE.  
RX MEDLINE-92037649; PubMed-1935979;  
RT Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;  
RT "limited proteolysis of tetanus toxin. Relation to activity and  
RT identification of cleavage sites";  
RL Eur. J. Biochem. 202:41-51(1991).  
RN [7]  
RP IDENTIFICATION AS ZINC-PROTEASE.  
RX MEDLINE-93010948; PubMed-1396558;  
RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,  
RA Montecucco C.;  
RT "Tetanus toxin is a zinc protein and its inhibition of  
RT neurotransmitter release and protease activity depend on zinc.";  
RL EMBO J. 11:3577-3583(1992).  
RN [8]  
RP IDENTIFICATION OF SUBSTRATE.  
RX MEDLINE-93063293; PubMed-1331807;  
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,  
RA Dasgupta B.R., Montecucco C.;  
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release  
RT by proteolytic cleavage of synaptobrevin.";  
RL Nature 359:832-835(1992).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.  
RX MEDLINE-97475217; PubMed-9334741;  
RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,  
RA Sax M.;  
RT "Structure of the receptor binding fragment HC of tetanus  
RT neurotoxin";  
RL Nat. Struct. Biol. 4:788-792(1997).  
RN [10]  
RP FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77  
CC BOND OF SYNAPTOBREVIN-2.  
CC [11]  
RP CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in  
CC synaptobrevin 2.  
CC [12]  
RP COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC [13]  
RP SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO  
CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE  
CC AND ARE NON-TOXIC AFTER SEPARATION.  
CC [14]  
RP MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO  
CC GLYGLIOSIDE RECEPTORS.  
CC [15]  
RP SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X04436; CAA28033.1; -;  
CC EMBL; X06214; CAA29564.1; -;  
CC EMBL; AF528097; AAO37454.1; -;  
CC EMBL; M12739; AAA23282.1; -;  
CC PIR; A25689; BTCLTN.  
CC PDB; 1AF9; 29-APR-98.  
CC PDB; 1A8D; 14-OCT-98.  
CC PDB; 1D0H; 27-MAR-00.  
CC PDB; 1DFQ; 24-MAR-00.  
CC PDB; 1DIW; 24-MAR-00.  
CC PDB; 1DLL; 24-MAR-00.  
CC PDB; 1FV3; 05-SEP-01.  
CC MEROPS; M27.001; -;  
CC InterPro; IPR000395; Bontoxilysin.  
CC InterPro; IPR006025; Zn\_Mtpeptidse.  
CC Pfam; PF01742; Peptidase\_M27; 1.

DR PRINTS; PR00760; BONTOXILYSIN.  
DR ProDom; PD001963; Bontoxilysin; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;  
KW 3D-structure; Complete proteome.  
FT INIT\_MET 0 0  
FT CHAIN 0 1 456 TETANUS TOXIN LIGHT CHAIN.  
FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.  
FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 233 233 BY SIMILARITY.  
FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).  
FT TRANSMEM 226 246 POTENTIAL.  
FT TRANSMEM 669 689 POTENTIAL.  
FT DISULFID 438 466 INTERCHAIN.  
FT DISULFID 1076 1092  
FT HELIX 876 882  
FT TURN 883 883  
FT STRAND 884 891  
FT TURN 892 893  
FT STRAND 894 897  
FT STRAND 904 907  
FT TURN 909 910  
FT STRAND 912 915  
FT STRAND 920 925  
FT TURN 928 929  
FT STRAND 932 935  
FT HELIX 938 940  
FT TURN 941 946  
FT STRAND 949 956  
FT HELIX 962 968  
FT TURN 969 970  
FT STRAND 972 977  
FT STRAND 980 981  
FT HELIX 983 985  
FT STRAND 987 995  
FT TURN 996 997  
FT STRAND 998 1004  
FT TURN 1006 1007  
FT STRAND 1010 1016  
FT STRAND 1020 1020  
FT TURN 1021 1022  
FT STRAND 1031 1037  
FT TURN 1039 1040  
FT STRAND 1042 1047  
FT TURN 1048 1049  
FT STRAND 1050 1056  
FT TURN 1058 1059  
FT STRAND 1068 1074  
FT TURN 1079 1080  
FT STRAND 1082 1091  
FT HELIX 1097 1105  
FT TURN 1106 1107  
FT STRAND 1112 1112  
FT STRAND 1114 1114  
FT TURN 1116 1117  
FT STRAND 1120 1120  
FT STRAND 1122 1122  
FT TURN 1123 1124  
FT STRAND 1127 1131  
FT HELIX 1132 1134  
FT TURN 1135 1136  
FT STRAND 1137 1141  
FT TURN 1144 1145  
FT STRAND 1148 1152  
FT STRAND 1155 1158  
FT TURN 1159 1162  
FT STRAND 1163 1166  
FT STRAND 1173 1178  
FT TURN 1184 1185  
FT STRAND 1188 1188  
FT STRAND 1190 1190  
FT TURN 1191 1192  
FT STRAND 1193 1201

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:00:31 : Search time 9.27978 Seconds  
(without alignments)  
253.383 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHWSYGLRPGSGPSLFNNF.....VSASHLEGPSLHWSYGLRPX 50

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	41.6	1314	1 TETX_CLOTE	P04958 clostridium
2	67	24.5	1296	1 BXG_CLOBO	Q60393 clostridium
3	65	23.7	1295	1 BXA1_CLOBO	P10845 legionella
4	63	23.0	431	1 PURA_LEGPN	Q8rm2 legionella
5	63	23.0	1274	1 BXF_CLOBO	P30996 clostridium
6	63	23.0	1290	1 BXB_CLOBO	P10844 clostridium
7	62.5	22.8	1290	1 BXC1_CLOBO	P18640 clostridium
8	62	22.6	1051	1 VP2_AHVS6	Q45894 clostridium
9	62	22.6	1295	1 BXA2_CLOBO	Q71024 african hor
10	61.5	22.4	92	1 GON1_RAT	P07490 rattus norv
11	60	21.9	92	1 GON1_TUPCB	Q95335 tupaja glis
12	59	21.5	67	1 GON1_MACMU	P35247 macaca mula
13	59	21.5	92	1 GON1_HUMAN	P01148 homo sapien
14	58.5	21.4	1250	1 BXE_CLOBO	Q00496 clostridium
15	58	21.2	61	1 GON1_SHEEP	Q28588 ovis aries
16	58	21.2	63	1 GON1_MESAU	Q09163 mesocricetu
17	58	21.2	89	1 GON1_XENLA	P45656 xenopus lae
18	58	21.2	90	1 GON1_MOUSE	P35622 mus musculu
19	58	21.2	91	1 GON1_PIG	P49921 sus scrofa
20	57.5	21.0	352	1 COA2_SV40	P03093 simian viru
21	57.5	21.0	842	1 ABC6_HUMAN	Q9np58 homo sapien
22	56.5	20.6	94	1 GON1_HAPHU	P51918 haplochromi
23	56.5	20.6	551	1 SVK_BRUME	Q8p2m8 brucella me
24	56.5	20.6	551	1 SVK_BRUSU	P39225 brucella su
25	56.5	20.6	658	1 STC1_STAAU	P07767 staphylococ
26	56.5	20.6	1250	1 BXE_CLOBO	P30995 clostridium
27	56	20.4	92	1 GON1_CHICK	P37042 gallus gall
28	56	20.4	141	1 AMPM_MYCCA	Q59509 mycoplasma
29	56	20.4	469	1 CG51_HUMAN	Q9p512 homo sapien
30	55.5	20.3	549	1 FXN_BRAJA	Q30373 bradyrhizob
31	55.5	20.3	715	1 STX2_STABU	P17855 staphylococ
32	54.5	19.9	464	1 VNSS_TSWY1	P26002 tomato spot
33	54.5	19.9	3122	1 DPOZ_MOUSE	Q61493 mus musculu

34	54.5	19.9	5217	1 HTS1_COCCA	Q01886 cochlilobolu
35	54	19.7	10	1 GON1_ALLMI	P37041 alligator m
36	54	19.7	585	1 YH70_SYNY3	P73627 synechocyst
37	53.5	19.5	634	1 PSTA_MYCGE	P47651 mycoplasma
38	53.5	19.5	760	1 AMY_CIOAB	P23671 clostridium
39	53	19.3	99	1 GON1_DICLA	Q91a10 dicentrarch
40	53	19.3	449	1 VNSS_INSVN	Q01811 impatiens n
41	53	19.3	746	1 PREA_PSEAE	Q05098 pseudomonas
42	53	19.3	1196	1 BXCN_CLOBO	P46081 clostridium
43	52.5	19.2	95	1 GON1_MORSA	O73812 morone saxa
44	52.5	19.2	584	1 GUND_CLOCE	P25472 clostridium
45	52	19.0	95	1 GON1_PAGMA	P70074 pagrus majo

## ALIGNMENTS

RESULT 1  
TETX\_CLOTE  
ID TETX\_CLOTE STANDARD; PRT: 1314 AA.  
AC P04958;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains:  
DE Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy  
DE chain (Tetanus toxin chain H)].  
GN TETX OR CTP60.  
OS Clostridium tetani.  
OG Plasmid pE88, and Plasmid 75 Kbp.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC PLASMID=75 Kbp;  
RX MEDLINE=87053814; PubMed=3536478;  
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,  
RA Weller U., Hudel M., Habermann E., Niemann H.,  
RT "Tetanus toxin: primary structure, expression in E. coli, and  
RT homology with botulinum toxins.";  
RL EMBO J. 5:2495-2502(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CN3911; PLASMID=75 Kbp;  
RX MEDLINE=87040747; PubMed=3774547;  
RA Fairweather N.F., Lyness V.A.;  
RT "The complete nucleotide sequence of tetanus toxin.";  
RL Nucleic Acids Res. 14:7809-7812(1986).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Massachusetts / E88; PLASMID=pE88;  
RX MEDLINE=22457253; PubMed=12552129;  
RA Brueggemann H., Baumeier S., Fricke W.F., Wiezer A., Liesegang H.,  
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,  
RA Gottschalk G.;  
RT "The genome sequence of Clostridium tetani, the causative agent of  
RT tetanus disease.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).  
RN [4]  
RP SEQUENCE OF 742-1314 FROM N.A.  
RC PLASMID=75 Kbp;  
RX MEDLINE=86085672; PubMed=3510187;  
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;  
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin  
RT fragment C in Escherichia coli.";  
RL J. Bacteriol. 165:21-27(1986).  
RN [5]  
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.  
RX MEDLINE=90201034; PubMed=2108021;  
RA Krieglstein K., Henschen A., Weller U., Habermann E.;  
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups  
RT in tetanus toxin.";



```

medaka-type gonadotropin-releasing hormone precursor - Japanese medaka
C:Species: Oryzias latipes (Japanese medaka)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 20-Jan-2003
C:Accession: JC7393
R:Okubo, K.; Amano, M.; Yoshiura, Y.; Suetake, H.; Aida, K.
Biochem. Biophys. Res. Commun. 276, 298-303, 2000
A:Title: A novel form of gonadotropin-releasing hormone in the medaka, Oryzias latipes
A:Reference number: JC7393
A:Contents: Brain
A:Accession: JC7393
A:Molecule type: mRNA
A:Residues: 1-91 <OKU>
A:Cross-references: DDBJ:AB041333
C:Comment: This protein plays the roles as a hypophysiotropic factor, and a physio-og
C:Genetics:
A:Gene: mdgnrh
C:Superfamily: gonadoliberin
C:Keywords: brain

Query Match 22.4%; Score 61.5; DB 2; Length 91;
Best Local Similarity 32.7%; Pred. No. 0.99; Indels 5; Gaps 1;
Matches 17; Conservative 7; Mismatches 23;

Qy 2 HWSYGURPGSGPSFNFTVSFWLRV-----PKVSASHLEGPSLHWSYGLR 48
|||:||| ||| :|: ||| ||| :|:
Db 23 HWSEGLSPGCKRELKYPNTLENQIRLLNSNTPCSDLSHLESSLAKIYRIK 74

RESULT 13
RRRTG

```

gonadoliberin precursor rat  
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing factor  
N:Contents: gonadoliberin; prolactin release-inhibiting factor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 18-Jun-1999  
C:Accession: A40147; B26173; A48410  
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.  
Mol. Endocrinol. 3, 1257-1262, 1989  
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic  
A:Reference number: A40147; MUID:89384661; PMID:2476669  
A:Accession: A40147  
A:Molecule\_type: DNA

A:Cross-references: GB:N31670; NID:g204447; PIDN:AAA1264.1; PID:g204448  
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and growth hormone releasing hormone  
A:Reference number: A94090; MUID:86094338; PMID:2867548  
A:Accession: B26173  
A:Molecule type: mRNA  
A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12579; NID:g204445; PIDN:AAA1263.1; PID:g204446  
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.  
Cell. Mol. Neurobiol. 12, 447-454, 1992  
A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone releasing hormone  
A:Reference number: A48410; MUID:93105480; PMID:1468115  
A:Accession: A48410  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-92 <MAI>  
A:Cross-references: GB:S50870; NID:g262059; PIDN:AAB34572.1; PID:g262060

A:Experimental source: thymus  
A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBIP:121083)  
C:Genetics:  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: stimulates pituitary secretion of lutropin and follitropin  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin

F:1-23/domain: signal sequence #status predicted <SIG>  
F:24-92/Product: progadoliblerin #status predicted <PGN>  
F:24-33/Product: gonadoliblerin #status predicted <GIN>

## RESULT 8

A48940  
bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum  
N:Alternate names: botulinum neurotoxin type B (BoNT/B)  
C:Species: Clostridium botulinum  
C>Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Jun-1999  
C:Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574  
R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.  
Appl. Environ. Microbiol. 58, 2345-2354, 1992  
A:Title: Molecular cloning of the Clostridium botulinum structural gene encoding the type B neurotoxin  
A:Reference number: A48940; MUID:92384550; PMID:1514783  
A:Accession: A48940  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1291 <WHE>  
A:Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735  
A:Experimental source: type B, Danish  
A:Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBI:P112081); this publication  
R:Campbell, K.D.; Collins, M.D.; East, A.K.  
J. Clin. Microbiol. 31, 2255-2262, 1993  
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific isoforms  
A:Reference number: S48103; MUID:94013372; PMID:8408542  
A:Accession: S48105  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 634-994 <CAM>  
A:Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783  
A:Experimental source: proteolytic type B, strain NCTC 7273  
R:Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.  
submitted to the EMBL Data Library, April 1992  
A:Description: Partial amino acid sequence of botulinum neurotoxin type B and comparison with other sequences  
A:Reference number: S21575  
A:Accession: S21575  
A:Molecule type: DNA  
A:Residues: 36-217, 'G', 219-224, 'S', 226-246 <SAZ>  
A:Cross-references: EMBL:Z11934; NID:g40383; PIDN:CAA77991.1; PID:g40384  
R:Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Gräbenstein, O.; Wernars, K.  
J. Biol. Chem. 267, 14721-14729, 1992  
A:Title: Minimal essential domains specifying toxicity of the light chains of tetanus toxin  
A:Reference number: A42871; MUID:92340509; PMID:1634516  
A:Accession: A42871  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-313, 'S', 315-451 <KUR>  
A:Experimental source: strain Okra  
A:Note: sequence extracted from NCBI backbone (NCBI:P109365)  
R:DasGupta, B.R.; Datta, A.  
Biochimie 70, 811-817, 1988  
A:Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with other sequences  
A:Reference number: S07155; MUID:89000987; PMID:3139097  
A:Accession: S07155  
A:Molecule type: protein  
A:Residues: 2-29, 'M', 31-45 <DAS>  
A:Accession: S08562  
A:Molecule type: protein  
A:Residues: 442-463, 'R', 465-467 <DA2>  
R:Schmidt, J.J.; Sathyamcorthy, V.; DasGupta, B.R.  
Arch. Biochem. Biophys. 238, 544-548, 1985  
A:Title: Partial amino acid sequences of botulinum neurotoxins types B and E.  
A:Reference number: S07128; MUID:85197963; PMID:3888113  
A:Accession: S07128  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 2-16 <SCH1>  
A:Accession: S08573  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 2-17 <SCH2>  
A:Accession: S08574  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 442-459 <SCH3>

R:Schiaivo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, Nature 359, 832-835, 1992  
A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolysis of syntaxin  
A:Reference number: S27125; MUID:93063293; PMID:1331807  
A:Contents: annotation  
C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synapses  
C:Genetics:  
A:Gene: bont/b  
C:Function:  
A:Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2  
C:Superfamily: tetanus toxin  
C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc  
F;2-441/Product: bontoxilysin B light chain #status experimental <LGHTR>  
F;442-1291/Product: bontoxilysin B heavy chain #status experimental <LGHTR>  
F;230,234/Binding site: zinc (His) #status predicted  
F;231/Active site: Glu #status predicted  
Query Match 23.08; Score 63; DB 1; Length 1291;  
Best Local Similarity 62.5%; Pred. No. 13;  
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
Qy 15 SLFNNFTVSFWLRVPK 30  
Db 921 SVFLDFSVFWIRPK 936  
RESULT 9  
S46431  
botulinum neurotoxin C1 - Clostridium botulinum phage 1C (strain C 468)  
N:Alternate names: BoNT/C1 protein  
C:Species: Clostridium botulinum phage 1C  
A:Variety: strain C 468  
C>Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 16-Jul-1999  
C:Accession: S46431; S49107  
R:Hauser, D.; Eklund, M.W.; Boquet, P.; Popoff, M.R.  
Mol. Gen. Genet. 243, 631-640, 1994  
A:Title: Organization of the botulinum neurotoxin C1 gene and its associated non-toxicity factor  
A:Reference number: S46426; MUID:94301293; PMID:8028579  
A:Accession: S46431  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1291 <HAU>  
A:Cross-references: EMBL:X72793; NID:g516171; PIDN:CAA51313.1; PID:g516175  
A:Experimental source: strain C 468  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993  
C:Superfamily: tetanus toxin  
Query Match 22.88; Score 62.5; DB 2; Length 1291;  
Best Local Similarity 25.4%; Pred. No. 15;  
Matches 17; Conservative 13; Mismatches 12; Indels 25; Gaps 3;  
Qy 5 YGLRPGSSGP-----SLENNFTVSFWLRVPKVSASHGPGSL----- 41  
Db 907 FDFKLGSSGDRGKVIYVTONENIVNYSFISFWIRINK-WWSNLPGYTIIDSVKNN 965  
Qy 42 -HWSYGL 47  
Db 966 SGWSIGI 972  
RESULT 10  
A49777  
botulinum neurotoxin type C1 precursor - Clostridium botulinum (type C, strain c-st)  
C:Species: Clostridium botulinum  
C>Date: 10-Mar-1994 #sequence\_revision 07-Apr-1994 #text\_change 17-Mar-2003  
C:Accession: S11291; A35396; S22166; A49777  
R:Hauser, D.; Eklund, M.W.; Kurazono, H.; Binz, T.; Niemann, H.; Gill, D.M.; Boquet, Nucleic Acids Res. 18, 4924, 1990  
A:Title: Nucleotide sequence of Clostridium botulinum C1 neurotoxin.  
A:Reference number: S11291; MUID:90370487; PMID:2204031  
A:Accession: S11291  
A:Status: preliminary  
A:Molecule type: DNA

QY 15 SLFNNFTVSFWLRVPK 30  
 I::||:||||:|:|  
 Db 936 SWYENFSVFWIRPK 951  
 RESULT 4  
 S33411  
 botulinum neurotoxin type F - Clostridium barati  
 C:Species: Clostridium barati  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 16-Jul-1999  
 C:Accession: S33411; S31860  
 R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T.  
 FEMS Microbiol. Lett. 108, 175-182, 1993  
 A:Title: Nucleotide sequence of the gene coding for Clostridium barati type F neurotoxin  
 A:Reference number: S33411; MUID:93252228; PMID:8486245  
 A:Accession: S33411  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1268 <THO>  
 A:Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139  
 C:Superfamily: tetanus toxin  
 C:Keywords: neurotoxin  
 Query Match 23.4%; Score 64; DB 2; Length 1268;  
 Best Local Similarity 62.5%; Pred. No. 9.6;  
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 15 SLFNNFTVSFWLRVPK 30  
 I::||:||||:|:|  
 Db 920 SRYQNFVSFWIRPK 935  
 RESULT 5  
 I40631  
 non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum  
 C:Species: Clostridium botulinum  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 16-Jul-1999  
 C:Accession: I40631; S48103; S48104; S36015  
 R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.  
 Curr. Microbiol. 28, 101-110, 1994  
 A:Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulinum  
 A:Reference number: I40631; MUID:94122659; PMID:7764370  
 A:Accession: I40631  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1291 <RES>  
 A:Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149  
 R:Campbell, K.D.; Collins, M.D.; East, A.K.  
 J. Clin. Microbiol. 31, 2255-2262, 1993  
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id  
 A:Reference number: S48103; MUID:94013372; PMID:8408542  
 A:Accession: S48103  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 634-761, 'B', 763-841, 'M', 843, 'T', 845, 'N', 847-994 <CAM1>  
 A:Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779  
 A:Experimental source: non-proteolytic strain 2129B (Scott)  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993  
 A:Accession: S48104  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 634-843, 'N', 845, 'N', 847-994 <CAM2>  
 A:Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781  
 A:Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)  
 C:Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic s  
 C:Genetics:  
 A:Gene: bont/b  
 C:Superfamily: tetanus toxin  
 C:Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc  
 F:2-441/Product: botulinum neurotoxin type B light chain #status predicted <LIGHT>  
 F:442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <H>  
 F:230,234/Binding site: zinc (His) #status predicted

F:231/Active site: Glu #status predicted

Query Match 23.4%; Score 64; DB 2; Length 1291;  
 Best Local Similarity 62.5%; Pred. No. 9.8;  
 Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30  
 I::||:||||:|:|  
 Db 921 SMFLDFSFWIRPK 936

#### RESULT 6

S48109

neurotoxin type F - Clostridium botulinum (fragment)

C:Species: Clostridium botulinum

C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 16-Jul-1999

C:Accession: S48109

R:Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A:Title: Gene probes for identification of the botulinum neurotoxin gene and specif

A:Reference number: S48103; MUID:94013372; PMID:8408542

A:Accession: S48109

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-369 <CAM>

A:Cross-references: EMBL:X70820; NID:g407790; PIDN:CAA50151.1; PID:g407791

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1998

C:Superfamily: tetanus toxin

Query Match 23.0%; Score 63; DB 2; Length 369;  
 Best Local Similarity 56.2%; Pred. No. 3.2;  
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30  
 I::||:||||:|:|  
 Db 295 SRYQNFVSFWIRPK 310

#### RESULT 7

I40813

neurotoxin type F - Clostridium botulinum

C:Species: Clostridium botulinum

C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 16-Jul-1999

C:Accession: I40813; S48108

R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson

FEMS Microbiol. Lett. 96, 225-230, 1992

A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.

A:Reference number: I40644

A:Accession: I40813

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1274 <RES>

A:Cross-references: GB:M2906; NID:g144866; PIDN:AAA23263.1; PID:g144867

R:Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A:Title: Gene probes for identification of the botulinum neurotoxin gene and specif

A:Reference number: S48103; MUID:94013372; PMID:8408542

A:Accession: S48108

A:Status: preliminary; translation not shown

A:Molecule type: DNA

A:Residues: 634-1002 <CAM>

A:Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789

C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

Query Match 23.0%; Score 63; DB 2; Length 1274;  
 Best Local Similarity 56.2%; Pred. No. 13;  
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30  
 I::||:||||:|:|  
 Db 928 SRYQNFVSFWIRPK 943



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:30:06 ; Search time 17:59 Seconds  
(without alignments)  
273.361 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHWSYGLRPGSGSPLENNF.....VSASHLEGPLSHWSYGLRPX 50

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	41.6	1315	1 BTCLTN	tentoxilysin (EC 3
2	67	24.5	1297	2 S3791	neurotoxin - Clost
3	65	23.7	1296	1 BTCLAB	bontoxilysin (EC 3
4	64	23.4	1268	2 S33411	botulinum neurotox
5	64	23.4	1291	2 I40631	non-proteolytic bo
6	63	23.0	369	2 S48109	neurotoxin type F
7	63	23.0	1274	2 I40813	neurotoxin type F
8	63	23.0	1291	1 A48940	bontoxilysin (EC 3
9	62.5	22.8	1291	2 S46431	botulinum neurotox
10	62.5	22.8	1291	2 A49777	botulinum neurotox
11	62	22.6	1296	2 I40645	botulinum neurotox
12	61.5	22.4	91	2 JC7393	medaka-type gonad
13	61.5	22.4	92	1 RHRTG	gonadoliberin prec
14	61	22.3	366	2 S48110	neurotoxin type F
15	61	22.3	502	2 T36589	probable transmemb
16	59.5	21.7	1285	2 S70582	botulinum neurotox
17	59.5	21.7	1999	2 AB2018	hypothetical prote
18	59	21.5	67	2 I78541	gonadoliberin prec
19	59	21.5	92	1 RHUG	gonadoliberin prec
20	58	21.2	10	1 RHFGG	gonadoliberin - pi
21	58	21.2	10	1 RHSHG	gonadoliberin - sh
22	58	21.2	89	2 I51423	gonadoliberin prec
23	58	21.2	90	1 RVMSG	gonadoliberin prec
24	57.5	21.0	352	1 VVVP24	coat protein VP2 -
25	57.5	21.0	836	2 J80248	ATP-binding casset
26	57	20.8	665	2 H83403	hypothetical prote
27	56.5	20.6	98	2 I50739	gonadotropin-relea
28	56.5	20.6	345	2 A58519	hypothetical 345 p
29	56.5	20.6	367	2 S48106	neurotoxin type E

30	56.5	20.6	444	2 T24077	hypothetical prote
31	56.5	20.6	551	2 AC3572	lysine-tRNA ligase
32	56.5	20.6	658	2 S06744	staphylocoagulase
33	56.5	20.6	1251	2 JH0256	botulinum neurotox
34	56.5	20.6	1252	2 S21178	botulinum neurotox
35	56	20.4	92	2 I50644	gonadoliberin I pr
36	56	20.4	141	2 S48593	probable methionyl
37	56	20.4	812	2 T01618	hypothetical prote
38	55.5	20.3	549	1 A47468	cytochrome-c oxida
39	55.5	20.3	658	2 A89786	staphylocoagulase
40	55.5	20.3	715	2 A41511	staphylocoagulase
41	55	20.1	496	2 T38197	probable myb-like
42	54.5	19.9	464	1 MNVUWC	nonstructural prot
43	54.5	19.9	3122	2 T17202	DNA-directed DNA p
44	54.5	19.9	5232	2 A45086	HC-toxin synthetas
45	54	19.7	10	1 RHAQ1	gonadoliberin I -

ALIGNMENTS

RESULT 1

BTCLTN tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani

N:Alternate names: tetanus neurotoxin

C:Species: Clostridium tetani

C>Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 03-Jun-2002

C:Accession: A25689; A25757; A25194; B25194; A60759; S69346; S09364

R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel,

EMBO J. 5, 2495-2502, 1986

A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with B

A:Reference number: A25689; MUID:87053814; PMID:3536478

A:Accession: A25689

A:Molecule type: DNA

A:Residues: 1-1315 <EIS>

A:CROSS-References: GB:X04436; MID:g40769; PIDN:CAA28033.1; PID:g40770

R:Fairweather, N.F.; Lyness, V.A.

Nucleic Acids Res. 14, 7809-7812, 1986

A:Title: The complete nucleotide sequence of tetanus toxin.

A:Reference number: A25757; MUID:87040747; PMID:3774547

A:Accession: A25757

A:Molecule type: DNA

A:Residues: 1-1315 <FAI>

A:CROSS-References: GB:X06214; MID:g40773; PIDN:CAA29564.1; PID:g40774

A:Experimental source: strain CN3911

R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.

J. Bacteriol. 165, 21-27, 1986

A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C i

A:Reference number: A25194; MUID:86085672; PMID:3510187

A:Accession: A25194

A:Molecule type: DNA

A:Residues: 743-1315 <PA2>

A:CROSS-References: GB:M12739; MID:g144920; PIDN:AAA23282.1; PID:g144921

A:Accession: B25194

A:Molecule type: protein

R:Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.

Infect. Immun. 57, 3588-3593, 1989

A:Title: Isolation, purification, and characterization of fragment B, the NH-2-termi

A:Reference number: A60759; MUID:90035436; PMID:2478476

A:Accession: A60759

A:Molecule type: protein

A:Residues: 461-475 <WAT>

R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.

J. Immunol. 142, 394-402, 1989

A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.

A:Reference number: J50098; MUID:89093918; PMID:2463305

A:Contents: annotation: epitope region

R:Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,

Nature 359, 832-835, 1992

A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by prot

A:Reference number: S27125; MUID:93063293; PMID:1331807

A:Contents: annotation



```

RA Fujinaga Y., Oguma K., Ohvama T.;
RT "Role of C-Terminal Region of HA-33 Component of Botulinum Toxin in
RT Hemagglutination.";
RL Biochem. Biophys. Res. Commun. 288:650-657(2001).
DR EMBL; AB061780; BAB71749.1; -.
DR InterPro; IPR000395; Bontoxilysin.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1291 AA; 148869 MW; 4A21DB35B8743CF8 CRC64;

Query Match 22.8%; Score 62.5; DB 2; Length 1291;
Best Local Similarity 25.4%; Pred. No. 29;
Matches 17; Conservative 13; Mismatches 12; Indels 25; Gaps 3;

Qy 5 YGLRPGSSGP-----SLFNFTVSFWLRVPKVSASHLEGPGL----- 41
Db 907 FDFKLGSGSGEDRGKVVIVTQENIVYNSWESFSISFWIRNK-WVSNLPGYTIIDSVKNN 965

Qy 42 -HWSYGL 47
Db 966 SGWSIGI 972

RESULT 14
QBUIQ7 PRELIMINARY; PRT; 91 AA.
AC Q8JIO7;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Medaka-type gonadotropin-releasing hormone.
GN MDGNRH.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HNI;
RX MEDLINE=22133319; PubMed=12137956;
RA Okubo K., Mitani H., Naruse K., Kondo M., Shima A., Tanaka M.,
RA Asakawa S., Shimizu N., Yoshiura Y., Aida K.;
RT "Structural characterization of GnRH loci in the medaka genome.";
RL Gene 293:181-189(2002).
DR EMBL; AB041333; BAB16303.1; -.
DR EMBL; AB041336; BAC06419.1; -.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
DR CHAIN 22 31
FT PROSITE 91 AA; 10307 MW; 8CD32BF97A8A027B CRC64;
SQ SEQUENCE 91 AA; 10293 MW; 8CD32BF97A8A027B CRC64;

Query Match 22.4%; Score 61.5; DB 13; Length 91;
Best Local Similarity 32.7%; Pred. No. 1.9;
Matches 17; Conservative 7; Mismatches 23; Indels 5; Gaps 1;

Qy 2 HWSYGLRPGSSGSLFNFTVSFWLRV-----PKVSASHLEGPGLHWSYGLR 48
Db 23 HWSFGLSPGGKRELKYPNTLENQIRLNSNAPCSDLSHLESSLAKIYRIK 74

Search completed: September 24, 2003, 17:47:22
Job time : 48.4765 secs

RA Fujinaga Y., Oguma K., Ohvama T.;
RT "Role of C-Terminal Region of HA-33 Component of Botulinum Toxin in
RT Hemagglutination.";
RL Biochem. Biophys. Res. Commun. 288:650-657(2001).
DR EMBL; AB061780; BAB71749.1; -.
DR InterPro; IPR000395; Bontoxilysin.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1291 AA; 148869 MW; 4A21DB35B8743CF8 CRC64;

Query Match 22.8%; Score 62.5; DB 2; Length 1291;
Best Local Similarity 25.4%; Pred. No. 29;
Matches 17; Conservative 13; Mismatches 12; Indels 25; Gaps 3;

Qy 5 YGLRPGSSGP-----SLFNFTVSFWLRVPKVSASHLEGPGL----- 41
Db 907 FDFKLGSGSGEDRGKVVIVTQENIVYNSWESFSISFWIRNK-WVSNLPGYTIIDSVKNN 965

Qy 42 -HWSYGL 47
Db 966 SGWSIGI 972

RESULT 14
QBUIQ7 PRELIMINARY; PRT; 91 AA.
AC Q8JIO7;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Medaka-type gonadotropin-releasing hormone.
GN MDGNRH.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HNI;
RX MEDLINE=22133319; PubMed=12137956;
RA Okubo K., Mitani H., Naruse K., Kondo M., Shima A., Tanaka M.,
RA Asakawa S., Shimizu N., Yoshiura Y., Aida K.;
RT "Structural characterization of GnRH loci in the medaka genome.";
RL Gene 293:181-189(2002).
DR EMBL; AB074499; BAC06421.1; -.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
DR CHAIN 22 31
FT PROSITE 91 AA; 10293 MW; 8CD32BF97A8A027B CRC64;
SQ SEQUENCE 91 AA; 10293 MW; 8CD32BF97A8A027B CRC64;

Query Match 22.4%; Score 61.5; DB 13; Length 91;
Best Local Similarity 32.7%; Pred. No. 1.9;
Matches 17; Conservative 7; Mismatches 23; Indels 5; Gaps 1;

Qy 2 HWSYGLRPGSSGSLFNFTVSFWLRV-----PKVSASHLEGPGLHWSYGLR 48
Db 23 HWSFGLSPGGKRELKYPNTLENQIRLNSNAPCSDLSHLESSLAKIYRIK 74

RESULT 15
Q9DGC8 PRELIMINARY; PRT; 91 AA.
AC Q9DGC8;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Prepro-gonadotropin-releasing hormone.
GN MDGNRH OR MFGNRH.

```

```

OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20462954; PubMed=11006121;
RA Okubo K., Amano M., Yoshiura Y., Suetake H., Aida K.;
RT "A Novel Form of Gonadotropin-Releasing Hormone in the Medaka, Oryzias
RT latipes.";
RL Biochem. Biophys. Res. Commun. 276:298-303(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22133319; PubMed=12137956;
RA Okubo K., Mitani H., Naruse K., Kondo M., Shima A., Tanaka M.,
RA Asakawa S., Shimizu N., Yoshiura Y., Aida K.;
RT "Structural characterization of GnRH loci in the medaka genome.";
RL Gene 293:181-189(2002).
DR EMBL; AB041333; BAB16303.1; -.
DR EMBL; AB041336; BAC06419.1; -.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
DR CHAIN 22 31
FT PROSITE 91 AA; 10307 MW; A00F2BED6FD6E0B5 CRC64;
SQ SEQUENCE 91 AA; 10307 MW; A00F2BED6FD6E0B5 CRC64;

Query Match 22.4%; Score 61.5; DB 13; Length 91;
Best Local Similarity 32.7%; Pred. No. 1.9;
Matches 17; Conservative 7; Mismatches 23; Indels 5; Gaps 1;

Qy 2 HWSYGLRPGSSGSLFNFTVSFWLRV-----PKVSASHLEGPGLHWSYGLR 48
Db 23 HWSFGLSPGGKRELKYPNTLENQIRLNSNAPCSDLSHLESSLAKIYRIK 74

Search completed: September 24, 2003, 17:47:22
Job time : 48.4765 secs

```

Query Match 23.4%; Score 64; DB 2; Length 1291;  
Best Local Similarity 62.5%; Pred. No. 18;  
Matches 10; Conservative 5; Mismatches 1; Indels

Query Match 23.4%; Score 64; DB 2; Length 1291;  
Best Local Similarity 62.5%; Pred. No. 18;  
Matches 10; Conservative 5; Mismatches 1; Indels

RESULT 12  
Q9ACP4  
ID Q9ACP4  
PRELIMINARY;

RESULT 12  
Q9ACP4  
ID Q9ACP4  
PRELIMINARY;

RNA	[1] _strand 1995/
RP	SEQUENCE FROM N.A.
RC	STRAIN=A3(2);
RX	MEDLINE=21996410; PubMed=12000953;
RA	Bentley S.D., Chater K.F., Cerdeno-Tarraga A.M., Challis G.L.,
RA	Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.
RA	Harper D., Bateman A., Brown S., Chanda G., Chen C.W., Collins M.,
RA	Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA	Huang C.H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA	Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA	Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.
RA	Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA	Hopwood D.A.;
RT	"Complete genome sequence of the model actinomycete <i>Streptomyces</i>
NT	<i>coelicolor</i> A3(2).";
RL	Nature 417:141-147(2002).
DR	EMBL: AL590464; CAC36800.1;

DR InterPro: IPR008530; YD.  
DR TIGRFAMs: TIGR01643; YD\_repeat\_2x; 11.  
KW plasmid; Complete proteome.  
SO SEQUENCE 2082 AA: 224419 MW: 38667800.312665DB CRC64:

Query Match	23.4%	Score 64;	DB 16;	Length 2082;
Best Local Similarity	44.28;	Pred. No. 32;		
Matches 19.	Conservative	5.	Mismatches	9.
				Indels 10.
				Gaps 3.

Db 270 WSAG---GSSGA-----FTW

Db 270 WSAG---GSSGA-----FTW

OS *Clostridium botulinum*.  
OC Bacteria: Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC *Clostridium*.  
OX NCBI\_taxid=1491;  
[1]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-Yoichi;  
RX MEDLINE=Y1534265; PubMed=11676492;  
RA Saqane Y., Koushichi H., Watanabe T., Sunagaawa H., Inoue K.,

```
Query Match      23.4%; Score 64; DB 2; Length 1291;
Best Local Similarity 62.5%; Pred. No. 18;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

```
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RA Shumlin Z., Dianliang L.;
RT "Cloning and sequence analysis of tetanus toxin gene.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389424; AAK72964.2; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR006025; Zn_MTPeptide.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR PRODOM; PD001963; BONTOXILYSIN; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON_TER 1
FT NON_TER 1310 1310
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;

Query Match 41.8%; Score 114; DB 2; Length 1310;
Best Local Similarity 95.3%; Pred. No. 6.1e-06;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPRKVSASHLE 37
Db 947 MFNNFTVSFWLRVPRKVSASHLE 968

RESULT 3
Q45848
ID Q45848 PRELIMINARY; PRT; 361 AA.
AC Q45848;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Botulinum neurotoxin type B (Fragment).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and
specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL; X70819; CAA50150.1; -.
DR HSSP; P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 42131 MW; A2E0FFFC81F9533D CRC64;

Query Match 23.4%; Score 64; DB 2; Length 361;
Best Local Similarity 62.5%; Pred. No. 4.3;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPRK 30
Db 288 SMFLDFSVFWIRIPK 303

RESULT 4
Q45846
ID Q45846 PRELIMINARY; PRT; 361 AA.
AC Q45846;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Botulinum neurotoxin type B (Fragment).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and
specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL; X70814; CAA50145.1; -.
DR HSSP; P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 42175 MW; 533EA98735CD98E1 CRC64;

Query Match 23.4%; Score 64; DB 2; Length 361;
Best Local Similarity 62.5%; Pred. No. 4.3;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPRK 30
Db 288 SMFLDFSVFWIRIPK 303

RESULT 5
Q9X708
ID Q9X708 PRELIMINARY; PRT; 441 AA.
AC Q9X708;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Botulinum neurotoxin type B (Fragment).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RA Lalli G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,
Schiaivo G.;
RT "Functional characterisation of tetanus and botulinum neurotoxins
binding domains.";
RL J. Cell Sci. 112:2715-2724(1999).
DR EMBL; AJ242628; CAB43706.1; -.
DR HSSP; P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1
FT NON_TER 441
SQ SEQUENCE 441 AA; 52772 MW; 721D0B468E8C95A4 CRC64;

Query Match 23.4%; Score 64; DB 2; Length 441;
Best Local Similarity 62.5%; Pred. No. 5.4;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPRK 30
Db 79 SMFLDFSVFWIRIPK 94

RESULT 6
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:29:11; Search time 48.4765 Seconds  
(without alignments)  
266.163 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHNSYGLRPGSSGSLFNNF.....VSASHLEGPSLHWSYGLRPX 50

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	41.6	451	2	O9LA13 clostridium
2	114	41.6	1310	2	Q93N27 clostridium
3	64	23.4	361	2	Q45848 clostridium
4	64	23.4	361	2	Q45846 clostridium
5	64	23.4	441	2	Q9X708 clostridium
6	64	23.4	1268	2	Q45851 clostridium
7	64	23.4	1291	2	Q9ZAJ8 clostridium
8	64	23.4	1291	2	Q93G71 clostridium
9	64	23.4	1291	2	Q933K0 clostridium
10	64	23.4	1291	2	Q98077 clostridium
11	64	23.4	1291	2	Q8GR96 clostridium
12	64	23.4	2082	16	Q9ACP4 streptomyce
13	62.5	22.8	1291	2	Q93HT3 clostridium
14	61.5	22.4	91	13	Q8J1Q7 oryzias lat
15	61.5	22.4	91	13	Q9DGC8 oryzias lat
16	61	22.3	228	12	Q9J2H3 macaca mula

17	61	22.3	502	16	Q9X8T8
18	61	22.3	1278	2	O57236
19	60.5	22.1	363	10	Q9FE71
20	59.5	21.7	657	10	Q9XE83
21	59.5	21.7	1285	2	O45967
22	59.5	21.7	1285	2	Q9LBR1
23	59.5	21.7	1999	16	Q8YWB9
24	58	21.2	90	13	Q90Y63
25	58	21.2	91	13	Q9PRH0
26	57.5	21.0	234	12	Q9W9A5
27	57.5	21.0	234	12	Q92837
28	57.5	21.0	352	12	Q98VM1
29	57.5	21.0	352	12	Q9W9F7
30	57.5	21.0	352	12	Q910V5
31	57.5	21.0	426	2	Q8KN97
32	57.5	21.0	766	4	Q96ME8
33	57.5	21.0	836	11	O70595
34	57.5	21.0	842	11	Q9DC29
35	57.5	21.0	896	4	Q9HA07
36	57	20.8	370	10	Q8W414
37	57	20.8	430	2	Q9XAV1
38	57	20.8	665	16	Q912G5
39	57	20.8	866	5	Q9VF20
40	57	20.8	1280	2	Q9ZAJ5
41	56.5	20.6	94	13	Q8JFY3
42	56.5	20.6	113	2	Q9KW76
43	56.5	20.6	140	2	Q9K200
44	56.5	20.6	140	2	Q9K2T0
45	56.5	20.6	167	2	Q9KW83

#### ALIGNMENTS

RESULT 1

ID O9LA13 PRELIMINARY; PRT; 451 AA.  
AC O9LA13;  
DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
DE Tetanus toxin (Fragment).  
OS Clostridium tetani.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=20886;  
RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;  
RT "Fragment C of tetanus toxin.";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF154828; AAF73267.1;  
DR HSSP; P04958; 1A8D.  
DR InterPro; IPR001064; Crystallin.  
DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
FT NON\_TER  
SQ SEQUENCE 451 AA; 51823 MW; 69A8C5F030E6CD8E CRC64;

Query Match 41.6%; Score 114; DB 2; Length 451;  
Best Local Similarity 95.5%; Pred. No. 1.8e-06;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 16 LFNNFTVSFWLRVPKVSASHLE 37  
:|||||  
Db 82 MFNNFTVSFWLRVPKVSASHLE 103

RESULT 2

ID Q93N27 PRELIMINARY; PRT; 1310 AA.  
AC Q93N27;  
DT 01-DEC-2001 (TremBLrel. 19, Created)



Best Local Similarity 75.0%; Pred. No. 5.2e-09;  
Matches 21; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30  
|||||  
Db 3 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 30  
|||||

Search completed: September 24, 2003, 17:40:02  
Job time : 36.036 secs

FT XX /note= "LHRH antigenic epitope AAY68566"

PN XX WO9966952-A1.

XX XX 29-DEC-1999.

PD XX 21-JUN-1999; 99WO-US13975.

XX XX 20-JUN-1998; 98US-0100412.

XX XX 20-JUN-1998; 98US-0100414.

XX XX (UNBI-) UNITED BIOMEDICAL INC.

PA XX Wang CY;

XX XX WPI; 2000-160562/14.

XX XX New peptide immunogen containing luteinising hormone-releasing hormone antigen site and helper T cell epitope, for e.g. contraception and treatment of cancer

XX XX Claim 9; Page 71; 102pp; English.

PS XX The present sequence represents a peptide immunogen comprising an

XX XX invasive domain immunostimulatory peptide of yersinia sp., a

CC XX synthetic helper T cell (Th) epitope and a target antigen, luteinising

CC XX hormone-releasing hormone (LHRH). The synthetic Th epitope is derived

CC XX from a structured synthetic antigen library (SSAL) designated SSAL1 Th1.

CC XX SSAL Th1 is modelled after a promiscuous epitope taken from the F protein

CC XX of the Measles virus. The peptide immunogens cause induction of a

CC XX specific immune response to LHRH which is involved in regulation of

CC XX spermatogenesis, ovulation, oestrus, sexual development and secretion

CC XX of sex hormones. Provision of a promiscuous T helper epitope (which is

CC XX functional in genetically diverse subjects) provides optimum

CC XX immunogenicity to the B cell epitopes of the target antigen and thus

CC XX high antibody titres against the target antigen. The peptide immunogens

CC XX of the invention are used to vaccinate against mammalian LHRH, for use

CC XX as (reversible) contraceptive; control of hormone-dependent tumours

CC XX (cancer of prostate or breast, also endometriosis); to prevent boar

CC XX taint (and improve meat quality) and for immunocastration.

XX XX Sequence 45 AA;

XX XX Query Match 67.5%; Score 108; DB 21; Length 45;

XX XX Best Local Similarity 71.4%; Pred. No. 5.7e-09;

XX XX Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

Qy 3 LSEIKGVIVHRLGVEGSPSLHWSYGLRP 30

Db 19 ISEIKGVIVHKIEGIGCE--HWSYGLRP 44

RESULT 15

AAAY91179

ID AAAY91179 standard; peptide; 31 AA.

XX XX AAAY91179;

AC XX

XX XX 22-MAY-2000 (first entry)

DT XX

XX XX Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:59.

DE XX

XX XX Promiscuous T-cell epitope; measles virus F protein; MVF;

KW XX hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

KW XX luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

KW XX somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;

KW XX foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

KW XX Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;

XX XX cholesteryl ester transport protein; anti-arteriosclerotic.

OS XX Chimeric - Measles virus.

OS XX Chimeric - Rattus sp.

XX XX WO9966957-A2.

XX XX 29-DEC-1999.

XX XX 21-JUN-1999; 99WO-US13975.

XX XX 20-JUN-1998; 98US-0100412.

XX XX (UNBI-) UNITED BIOMEDICAL INC.

XX XX Wang CY;

XX XX WPI; 2000-160564/14.

XX XX New artificial T helper cell epitope and derived immunogens with target antigenic site, for immunization against e.g. malaria, arteriosclerosis or human immune deficiency virus

XX XX Example 1; Page 86; 129pp; English.

XX XX The invention relates to novel promiscuous T helper cell epitopes (Th), and immunogenic peptides comprising the Th epitopes and peptide immunogens along with B cell epitopes. The Th epitopes and peptide immunogens containing them, are used to induce a T helper cell response, specifically against Plasmodium falciparum, cholesteryl ester transport protein (CETP) or HIV epitopes, but more generally against any pathogen, immunoreactive self-antigen or tumour antigen. The Th epitopes and peptide immunogens may be used for prevention and/or treatment of infections (HIV, foot-and-mouth disease or malaria); for cancer immunotherapy; for inhibition of the action of luteinising hormone releasing hormone (LHRH) for contraception, treatment of hormone-dependent cancer, prevention of boar taint in meat, and immunocastration; for promoting the growth of animals; or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence AAAY91121 represents a promiscuous T helper epitope from the measles virus F (MVF) protein and sequences AAAY91122-Y91142, AAAY91226 and AAAY91245-Y91246 represent synthetic Th epitopes based on the MVF Th epitope. Sequence AAAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences AAAY91144-Y91155 are synthetic epitopes derived from this HBV epitope. AAAY91156-Y91196, AAAY91227 and AAAY91242-Y91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. AAAY91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. AAAY91200 is somatostatin, and AAAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. AAAY91208 is a human CD4 CDR2-like domain antigenic site, and AAAY91209-Y90211 are MVF Th epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV infection of T cells. AAAY90212 is a modified version of a human IGE (immunoglobulin E) CH3 domain, and AAAY90213-Y90219 are Th epitope/IgE CH3 antigenic peptides which may be used in the treatment of allergies. AAAY91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and AAAY91221-Y91222 comprise this peptide and a Th epitope. AAAY91223 is a Plasmodium falciparum circumsporozoite (CS) target antigen, and AAAY91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be used in a malaria vaccine. AAAY91228-Y91231 represent CPTP-derived peptides and AAAY91232-Y91241 are immunogens comprising a CPTP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. AAAY91247 and AAAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAAY91248-Y91251 and AAAY91258-Y91273 are antigenic peptides comprising MVF Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. AAAY91198 and AAAY91199 are respectively an immunostimulatory invasive protein epitope from Yersinia species, and hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.

XX XX Sequence 31 AA;

XX XX Query Match 56.9%; Score 107; DB 21; Length 31;

```

SQ Sequence 27 AA;
Query Match 67.5%; Score 108; DB 21; Length 27;
Best Local Similarity 75.0%; Pred. No. 3.le-09;
Matches 21; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVVHRLEGVEGSPSLHWSYGLRP 30
:|||||||:||||:| |||||
Db 1 MSEIKGVVHKLEGGGE--HWSYGLRP 26

RESULT 13
AAAY91165
ID AAY91165 standard; peptide; 45 AA.
AC AAY91165;
XX
XX
XX
XX
XX 22-MAY-2000 (first entry)
DT
DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:45.
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesterol ester transport protein; anti-arteriosclerotic.
XX
OS Chimeric - Measles virus.
OS Chimeric - Rattus sp.
XX
XX
XX WO9966957-A2.
XX
XX
XX 29-DEC-1999.
XX
XX
XX 21-JUN-1999; 99WO-US13975.
XX
XX 20-JUN-1998; 98US-0100412.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX
XX Wang CY;
XX
XX WPI; 2000-160564/14.
XX
XX
XX New artificial T helper cell epitope and derived immunogens with target
XX antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX or human immune deficiency virus
XX
XX Example 1; Page 80; 129pp; English.
XX
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
XX and immunogenic peptides comprising the Th epitopes of the invention
XX along with B cell epitopes. The Th epitopes and peptide immunogens
XX containing them, are used to induce a T helper cell response.
XX specifically against Plasmodium falciparum, cholesterol ester transport
XX protein (CETP) or HIV epitopes, but more generally against any pathogen,
XX immunoreactive self-antigen or tumour antigen. The Th epitopes and
XX peptide immunogens may be used for prevention and/or treatment of
XX infections (HIV, foot-and-mouth disease or malaria); for cancer
XX immunotherapy; for inhibition of the action of luteinising hormone
XX releasing hormone (LHRH) for contraception, treatment of hormone-
XX dependent cancer, prevention of boar taint in meat, and
XX immunocastration; for promoting the growth of animals; or for
XX treating allergies or arteriosclerosis. Incorporation of a promiscuous
XX Th (functional in genetically diverse subjects) into an immunogen
XX improves capacity to induce a strong T helper cell-mediated immune
XX response, resulting in production of antibodies against a target
XX antigen. Th can replace carrier proteins and pathogen-derived T helper
XX epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
XX from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
XX AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the

```

---

```

CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human IGE
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CETP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91138 and AAY91199 are respectively an immunostimulatory invasive
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.
XX
XX
XX Query Match 67.5%; Score 108; DB 21; Length 45;
XX Best Local Similarity 71.4%; Pred. No. 5.7e-09;
XX Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVVHRLEGVEGSPSLHWSYGLRP 30
:|||||||:||||:| |||||
Db 19 ISEIKGVVHKIEGIGGE--HWSYGLRP 44

RESULT 14
AAAY68573
ID AAY68573 standard; peptide; 45 AA.
XX
XX
XX AAY68573;
XX
XX
XX 05-MAY-2000 (first entry)
DT
DE Peptide immunogen comprising a Th epitope and LHRH target antigen.
XX
XX Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;
XX luteinising hormone-releasing hormone; spermatogenesis; ovulation;
XX oestrus; sexual development; sex hormone; promiscuous T helper epitope;
XX vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
XX breast cancer; endometriosis; boar taint; meat quality;
XX invasin domain; immunocastration.
XX
XX Synthetic.
OS Yersinia sp.
OS Measles virus.
OS Unidentified.
XX
XX Key Location/Qualifiers
XX Peptide 1..16
XX /note= "invasin domain AAY68565"
XX Peptide 17..18
XX /note= "spacer"
XX Peptide 19..33
XX /note= "helper Th epitope AAY68544"
XX Peptide 34..35
XX /note= "spacer"
XX Peptide 36..45

```

CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
 CC immunotherapy; for inhibition of the action of luteinising hormone  
 CC releasing hormone (LHRH) for contraception, treatment of hormone-  
 CC dependent cancer, prevention of boar taint in meat, and  
 CC immunocastration); for promoting the growth of animals; or for  
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous  
 CC Th (functional in genetically diverse subjects) into an immunogen  
 CC improves capacity to induce a strong T helper cell-mediated immune  
 CC response, resulting in production of antibodies against a target  
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper  
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope  
 CC from the measles virus F (MV) protein and sequences AAY91122-Y91142,  
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the  
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope  
 CC from hepatitis B virus (HBV) surface antigen, and sequences  
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.  
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides  
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197  
 CC is the LHRH target antigenic peptide used in these LHRH antigenic  
 CC peptides comprising somatostatin and a Th epitope. Somatostatin  
 CC peptides may be used to promote growth in livestock. AAY91208 is a  
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th  
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV  
 CC infection of T cells. AAY90212 is a modified version of a human IGE  
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3  
 CC antigenic peptides which may be used in the treatment of allergies.  
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)  
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th  
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target  
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th  
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent  
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a  
 CC CERP peptide and a Th epitope which may be used to prevent or treat  
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257  
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and  
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell  
 CC epitopes which may be used as a component in an anti-HIV-1 vaccine.  
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive  
 CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of  
 CC which may optionally be used in the antigenic peptides of the  
 CC invention.

Qy Sequence 27 AA;  
 Query Match 67.5%; Score 108; DB 21; Length 27;  
 Best Local Similarity 71.4%; Pred. No. 3.1e-09;  
 Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;  
 3 LSEIKGVIVHLEGEVPSLHWSYGLRP 30  
 :|||||||:|:|:| |  
 1 ISEIKGVIVHLEGEVPSLHWSYGLRP 26

RESULT 12  
 AAY91167  
 ID AAY91167 standard; peptide; 27 AA.  
 XX  
 AC AAY911167;  
 XX  
 DT 22-MAY-2000 (first entry)  
 XX  
 DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:47.  
 XX  
 KW Promiscuous T-cell epitope; measles virus F protein; MVF;  
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;  
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
 KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;  
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;  
 KW cholesteryl ester transport protein; anti-arteriosclerotic.  
 XX  
 OS Chimeric - Measles virus.

OS Chimeric - Rattus sp.  
 XX WO9966957-A2.  
 PN  
 XX 29-DEC-1999.  
 PD  
 XX 21-JUN-1999; 99WO-US13975.  
 PF  
 XX 20-JUN-1998; 98US-0100412.  
 PR  
 XX (UNBI-) UNITED BIOMEDICAL INC.  
 PA  
 XX Wang CY;  
 PI  
 XX WPI; 2000-160564/14.  
 DR  
 XX New artificial T helper cell epitope and derived immunogens with target  
 XX antigenic site, for immunization against e.g. malaria, arteriosclerosis  
 XX or human immune deficiency virus -  
 PT  
 XX Example 1; Page 81; 129pp; English.  
 PS  
 XX The invention relates to novel promiscuous T helper cell epitopes (Th),  
 CC and immunogenic peptides comprising the Th epitopes and peptide immunogens  
 CC along with B cell epitopes. The Th epitopes and peptide immunogens  
 CC containing them, are used to induce a T helper cell response,  
 CC specifically against Plasmodium falciparum, cholesteryl ester transport  
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,  
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and  
 CC peptide immunogens may be used for prevention and/or treatment of  
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
 CC immunotherapy; for inhibition of the action of luteinising hormone  
 CC releasing hormone (LHRH) for contraception, treatment of hormone-  
 CC dependent cancer, prevention of boar taint in meat, and  
 CC immunocastration); for promoting the growth of animals; or for  
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous  
 CC Th (functional in genetically diverse subjects) into an immunogen  
 CC improves capacity to induce a strong T helper cell-mediated immune  
 CC response, resulting in production of antibodies against a target  
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper  
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope  
 CC from the measles virus F (MV) protein and sequences AAY91122-Y91142,  
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the  
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope  
 CC from hepatitis B virus (HBV) surface antigen, and sequences  
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.  
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides  
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197  
 CC is the LHRH target antigenic peptide used in these LHRH antigenic  
 CC peptides comprising somatostatin and a Th epitope. Somatostatin  
 CC immunogens may be used to promote growth in livestock. AAY91208 is a  
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th  
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV  
 CC infection of T cells. AAY90212 is a modified version of a human IGE  
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3  
 CC antigenic peptides which may be used in the treatment of allergies.  
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)  
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th  
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target  
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th  
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent  
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a  
 CC CERP peptide and a Th epitope which may be used to prevent or treat  
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257  
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and  
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell  
 CC epitopes which may be used as a component in an anti-HIV-1 vaccine.  
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive  
 CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of  
 CC which may optionally be used in the antigenic peptides of the  
 CC invention.

DT 22-MAY-2000 (first entry)

XX Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:55.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;

KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;

KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

KW Plasmodium falciparum; circumporoite; antimalarial; CPTP;

KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Chimeric - Measles virus.

OS Chimeric - Rattus sp.

XX WO9966957-A2.

XX 29-DEC-1999.

PF 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

PR (UNBI-) UNITED BIOMEDICAL INC.

PA Wang CY;

PI WPI; 2000-160564/14.

DR New artificial T helper cell epitope and derived immunogens with target

PT antigenic site, for immunization against e.g. malaria, arteriosclerosis

PT or human immune deficiency virus

XX Example 1; Page 84; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),

CC and immunogenic peptides comprising the Th epitopes and peptide immunogens

CC along with B cell epitopes. The Th epitopes and peptide immunogens

CC containing them, are used to induce a T helper cell response,

CC specifically against Plasmodium falciparum, cholesteryl ester transport

CC protein (CPTP) or HIV epitopes, but more generally against any pathogen,

CC immunoreactive self-antigen or tumour antigen. The Th epitopes and

CC peptide immunogens may be used for prevention and/or treatment of

CC infections (HIV, foot-and-mouth disease or malaria); for cancer

CC immunotherapy; for inhibition of the action of luteinising hormone

CC releasing hormone (LHRH) for contraception, treatment of hormone-

CC dependent cancer, prevention of boar taint in meat, and

CC immunocastration; for promoting the growth of animals; or for

CC treating allergies or arteriosclerosis. Incorporation of a promiscuous

CC Th (functional in genetically diverse subjects) into an immunogen

CC improves capacity to induce a strong T helper cell-mediated immune

CC response, resulting in production of antibodies against a target

CC antigen. Th can replace carrier proteins and pathogen-derived T helper

CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope

CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,

CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the

CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope

CC from hepatitis B virus (HBV) surface antigen, and sequences

CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.

CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides

CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197

CC is the LHRH target antigenic peptide used in these LHRH antigenic

CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic

CC peptides comprising somatostatin and a Th epitope. Somatostatin

CC immunogens may be used to promote growth in livestock...AAY91208 is a

CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th

CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV

CC infection of T cells. AAY90212 is a modified version of a human IgE

CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3

CC antigenic peptides which may be used in the treatment of allergies

CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)

CC VPI capsid protein and AAY91221-Y91222 comprise this peptide and a Th

CC epitope. AAY91223 is a Plasmodium falciparum circumporoite (CS) target

CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th

CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent

CC CPTP-derived peptides and AAY91232-Y91241 are immunogens comprising a

CC CPTP peptide and a Th epitope which may be used to prevent or treat

CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257

CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and

CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell

CC epitope which may be used as a component in an anti-HIV-1 vaccine.

CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin

CC protein epitope from Yersinia species, and hinge spacer peptide, both of

CC which may optionally be used in the antigenic peptides of the

CC invention.

XX Query Match 58.8%; Score 110; DB 21; Length 31;

XX Best Local Similarity 78.6%; Pred. No. 1.8e-09;

XX Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGVGSPSLHWSYGLRP 30

DB 3 LSEIKGVIVHKLEGVLFGEHWSYGLRP 30

RESULT 11

AAAY91161

ID AAY91161 standard; peptide; 27 AA.

XX AC AAY91161;

XX 22-MAY-2000 (first entry)

XX Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:41.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;

KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;

KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

KW Plasmodium falciparum; circumporoite; antimalarial; CPTP;

KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Chimeric - Measles virus.

OS Chimeric - Rattus sp.

XX WO9966957-A2.

XX 29-DEC-1999.

PF 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

PR (UNBI-) UNITED BIOMEDICAL INC.

PA Wang CY;

PI WPI; 2000-160564/14.

DR New artificial T helper cell epitope and derived immunogens with target

PT antigenic site, for immunization against e.g. malaria, arteriosclerosis

PT or human immune deficiency virus

XX Example 1; Page 79; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),

CC and immunogenic peptides comprising the Th epitopes of the invention

CC along with B cell epitopes. The Th epitopes and peptide immunogens

CC containing them, are used to induce a T helper cell response,

CC specifically against Plasmodium falciparum, cholesteryl ester transport

CC protein (CPTP) or HIV epitopes, but more generally against any pathogen,

CC immunoreactive self-antigen or tumour antigen. The Th epitopes and

CC peptide immunogens may be used for prevention and/or treatment of

CC infections (HIV, foot-and-mouth disease or malaria); for cancer

CC immunotherapy; for inhibition of the action of luteinising hormone

CC releasing hormone (LHRH) for contraception, treatment of hormone-

CC dependent cancer, prevention of boar taint in meat, and

CC immunocastration; for promoting the growth of animals; or for

CC treating allergies or arteriosclerosis. Incorporation of a promiscuous

CC Th (functional in genetically diverse subjects) into an immunogen

CC improves capacity to induce a strong T helper cell-mediated immune

CC response, resulting in production of antibodies against a target

CC antigen. Th can replace carrier proteins and pathogen-derived T helper

CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope

CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,

CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the

CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope

CC from hepatitis B virus (HBV) surface antigen, and sequences

CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.

CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides

CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197

CC is the LHRH target antigenic peptide used in these LHRH antigenic

CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic

CC peptides comprising somatostatin and a Th epitope. Somatostatin

CC immunogens may be used to promote growth in livestock...AAY91208 is a

CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th

CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV

CC infection of T cells. AAY90212 is a modified version of a human IgE

CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3

CC antigenic peptides which may be used in the treatment of allergies

CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)

CC VPI capsid protein and AAY91221-Y91222 comprise this peptide and a Th

CC epitope. AAY91223 is a Plasmodium falciparum circumporoite (CS) target

```

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C. H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ. ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-13

Query Match          76.0%; Score 112.5; DB 2; Length 27;
Best Local Similarity 85.2%; Pred. No. 5.2e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

OY      1 QYIKANSKFIGITELGSLHMSYGLRP 27
Db      3 QYIKANSKFIGITELG--HMSYGLRP 26

RESULT 5
US-08-446-692-57
; Sequence 57, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Zamb, Chang Yi
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Maria C. H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C. H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ. ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-57
```

```

Query Match          54.1%; Score 80; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 9.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 QYIKANSKFIGITELG 16
Db      5 QYIKANSKFIGITELG 20

RESULT 6
US-08-446-692-63
; Sequence 63, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Zamb, Chang Yi
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Maria C. H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C. H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ. ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-63

RESULT 7
US-08-488-351A-57
; Sequence 57, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Zamb, Chang Yi
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESSES:
```

ADDRESSEE: Maria C.H. Lin  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: US  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,351A  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/446,692  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/229,275  
FILING DATE: 14-APR-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/057,166  
FILING DATE: 27-APR-1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria C.H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4146 US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)415-8745  
TELEFAX: (516)751-6849  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-488-351A-57

Query Match 54.1%; Score 80; DB 2; Length 37;  
Best Local Similarity 100.0%; Pred. No. 9.4e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYKANSKFQITELG 16  
DB 21 QYKANSKFQITELG 36

RESULT 8  
US-08-488-351A-63  
Sequence 63, Application US/08488351A  
Patent No. 5843446  
GENERAL INFORMATION:  
APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang Yi  
APPLICANT: Zamb, Timothy  
TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maria C.H. Lin  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: US  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,351A  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/446,692  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/229,275  
FILING DATE: 14-APR-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/057,166  
FILING DATE: 27-APR-1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria C.H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4146 US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)415-8745  
TELEFAX: (516)751-6849  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-488-351A-63

Query Match 54.1%; Score 80; DB 2; Length 37;  
Best Local Similarity 100.0%; Pred. No. 9.4e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYKANSKFQITELG 16  
DB 5 QYKANSKFQITELG 20

RESULT 9  
US-08-945-289-8  
Sequence 8, Application US/08945289  
Patent No. 6555113  
GENERAL INFORMATION:  
APPLICANT: Rittershaus, Charles, W.  
APPLICANT: Thomas, Lawrence J.  
TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER  
TITLE OF INVENTION: TRANSFER PROTEIN (CETP) ACTIVITY  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Yankwich & Associates  
STREET: 130 Bishop Allen Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02139  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/945,289  
FILING DATE: October 17, 1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/432,483  
FILING DATE: May 1, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Leon R. Yankwich  
REGISTRATION NUMBER: 30,237

```

; REFERENCE/DOCKET NUMBER: TCS-411.1P US
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL:
; ANTI-SENSE:
; FEATURE:
; NAME/KEY:
; LOCATION:
; US-08-945-289-8

Query Match          52.7%; Score 78; DB 4; Length 50;
Best Local Similarity 94.1%; Pred. No. 2.7e-05;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYKANSKFIGITELGP 17
   |||||||||
Db 2 QYKANSKFIGITELFP 18

RESULT 10
US-08-464-496-25
; Sequence 25, Application US/08464496
; Patent No. 6322789
; GENERAL INFORMATION:
; APPLICANT: EpiImmune, Inc.
; APPLICANT: Vitello, Maria
; APPLICANT: Chesnut, Robert
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; FILE REFERENCE: 39963-20001.13
; CURRENT APPLICATION NUMBER: US/08/464,496
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: 07/935,811
; PRIOR FILING DATE: 1992-08-26
; PRIOR APPLICATION NUMBER: 07/874,491
; PRIOR FILING DATE: 1992-04-27
; PRIOR APPLICATION NUMBER: 07/827,682
; PRIOR FILING DATE: 1992-01-29
; PRIOR APPLICATION NUMBER: 07/749,568
; PRIOR FILING DATE: 1991-08-26
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Oligopeptide
; US-08-464-496-25

Query Match          52.0%; Score 77; DB 4; Length 24;
Best Local Similarity 88.9%; Pred. No. 1.7e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYKANSKFIGITELGPS 18
   |||||||||
Db 1 QYKANSKFIGITELFPS 18

RESULT 11
US-08-197-484-110
; Sequence 110, Application US/08197484
; Patent No. 6419931
; GENERAL INFORMATION:
; APPLICANT: VITELLO, Maria A.
; APPLICANT: CHESTNUT, Robert W.
; APPLICANT: SETTE, Alessandro D.
; APPLICANT: CELIS, Esteban
; APPLICANT: GRAY, Howard
```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,484
; FILING DATE: 16-FEB-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-197-484-110

Query Match          52.0%; Score 77; DB 4; Length 24;
Best Local Similarity 88.9%; Pred. No. 1.7e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYKANSKFIGITELGPS 18
   |||||||||
Db 1 QYKANSKFIGITELFPS 18

RESULT 12
PCT-US92-07218-25
; Sequence 25, Application PC/TUS9207218
; GENERAL INFORMATION:
; APPLICANT: Vitello, Maria A.
; APPLICANT: Chesnut, Robert W.
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPTOPES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Stewart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/07218  
FILING DATE: 19920826  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,568  
FILING DATE: 26-AUG-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,682  
FILING DATE: 29-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 14137-26-3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US92-07218-25

Query Match 52.0%; Score 77; DB 5; Length 24;  
Best Local Similarity 88.9%; Pred. No. 1.7e-05;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 OYKANSKFIGITELGPS 18  
Db 1 OYKANSKFIGITELGPS 18

RESULT 13  
PCT-US95-02121-110  
Sequence 110, Application PC/TUS9502121  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
TITLE OF INVENTION: CTL IMMUNITY  
NUMBER OF SEQUENCES: 153  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02121  
FILING DATE: 16-FEB-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/197,484  
FILING DATE: 16-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/935,811  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,682  
FILING DATE: 29-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,568

FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 14137-26-4PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 110:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
PCT-US95-02121-110

Query Match 52.0%; Score 77; DB 5; Length 24;  
Best Local Similarity 88.9%; Pred. No. 1.7e-05;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 OYKANSKFIGITELGPS 18  
Db 1 OYKANSKFIGITELGPS 18

RESULT 14  
US-08-464-496-32  
Sequence 32, Application US/08464496  
Patent No. 6322789  
GENERAL INFORMATION:  
APPLICANT: Epimmune, Inc.  
APPLICANT: Vitello, Maria  
TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL  
TITLE OF INVENTION: EPITOPES  
FILE REFERENCE: 39963-20001.13  
CURRENT APPLICATION NUMBER: US/08/464,496  
CURRENT FILING DATE: 1995-06-05  
PRIOR APPLICATION NUMBER: 07/935,811  
PRIOR FILING DATE: 1992-08-26  
PRIOR APPLICATION NUMBER: 07/874,491  
PRIOR FILING DATE: 1992-04-27  
PRIOR APPLICATION NUMBER: 07/827,682  
PRIOR FILING DATE: 1992-01-29  
PRIOR APPLICATION NUMBER: 07/749,568  
PRIOR FILING DATE: 1991-08-26  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 32  
LENGTH: 27  
TYPE: PPT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Oligopeptide  
US-08-464-496-32

Query Match 52.0%; Score 77; DB 4; Length 27;  
Best Local Similarity 88.9%; Pred. No. 2e-05;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 OYKANSKFIGITELGPS 18  
Db 4 OYKANSKFIGITELGPS 21

RESULT 15  
PCT-US92-07218-32  
Sequence 32, Application PC/TUS9207218  
GENERAL INFORMATION:  
APPLICANT: Vitello, Maria A.  
APPLICANT: Chesnut, Robert W.  
TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL

1 TITLE OF INVENTION: EPITOPES  
 2 NUMBER OF SEQUENCES: 35  
 3 CORRESPONDENCE ADDRESS:  
 4 ADDRESSEE: Townsend and Townsend  
 5 STREET: One Market Plaza, Stewart Street Tower  
 6 CITY: San Francisco  
 7 STATE: California  
 8 COUNTRY: USA  
 9 ZIP: 94105  
 10  
 11 COMPUTER READABLE FORM:  
 12 MEDIUM TYPE: floppy disk  
 13 COMPUTER: IBM PC compatible  
 14 OPERATING SYSTEM: PC-DOS/MS-DOS  
 15 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 16 CURRENT APPLICATION DATA:  
 17 APPLICATION NUMBER: PCT/US92/07218  
 18 FILING DATE: 19920826  
 19 CLASSIFICATION: 424  
 20 PRIOR APPLICATION DATA:  
 21 APPLICATION NUMBER: US 07/749,568  
 22 FILING DATE: 26-AUG-1991  
 23 PRIOR APPLICATION DATA:  
 24 APPLICATION NUMBER: US 07/827,682  
 25 FILING DATE: 29-JAN-1992  
 26 PRIOR APPLICATION DATA:  
 27 APPLICATION NUMBER: US 07/874,491  
 28 FILING DATE: 27-APR-1992  
 29 ATTORNEY/AGENT INFORMATION:  
 30 NAME: Smith, William M.  
 31 REGISTRATION NUMBER: 30,223  
 32 REFERENCE/DOCKET NUMBER: 14137-26-3  
 33 TELECOMMUNICATION INFORMATION:  
 34 TELEPHONE: 415-326-2400  
 35 TELEFAX: 415-543-5043  
 36 INFORMATION FOR SEQ ID NO: 32:  
 37 SEQUENCE CHARACTERISTICS:  
 38 LENGTH: 27 amino acids  
 39 TYPE: AMINO ACID  
 40 STRANDEDNESS: single  
 41 TOPOLOGY: linear  
 42 MOLECULE TYPE: protein  
 43  
 44 PCT-US92-07218-32

	Query Match	52.0%;	Score 77;	DB 5;	Length 27;
	Best Local Similarity	88.9%;	Pred. No. 2e-05;		
Matches	16;	Conservative	0;	Mismatches	2;
				Indels	0;
				Gaps	0;
Oy	1 OYKANSFCIGITELGPS	18			
		11			
db	4 OYKANSKFICGITTELP	21			

Search completed: September 24, 2003, 17:00:22  
Job time : 8.39917 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:57:10 : Search time 14.0388 Seconds  
(without alignments)  
301.778 Million cell updates/sec

Title: US-09-848-834a-11

Perfect score: 148  
Sequence: 1 OYIRANSKFITETELGPELSHWSYGLRPX 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA.\*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	99.3	28	9	US-09-848-834A-11
2	136	91.9	46	9	US-09-848-834A-19
3	112.5	76.0	27	12	US-10-076-674-7
4	81.5	55.1	31	12	US-10-237-656-1
5	81.5	55.1	33	9	US-09-848-834A-12
6	80	54.1	32	12	US-10-237-656-5
7	80	54.1	137	12	US-10-237-656-3
8	78	52.7	50	9	US-09-943-548-8
9	78	52.7	50	15	US-10-339-522-8
10	77	52.0	24	15	US-10-128-711-110
11	76.5	51.7	51	9	US-09-848-834A-20
12	76	51.4	31	9	US-09-943-548-2
13	76	51.4	31	15	US-10-339-522-2
14	74	50.0	15	9	US-09-862-849-2
15	74	50.0	15	10	US-09-785-215-4

16	74	50.0	15	12	US-10-237-656-13	Sequence 13, Appl
17	74	50.0	15	12	US-09-405-986-1	Sequence 1, Appl
18	74	50.0	15	12	US-10-223-809A-4	Sequence 4, Appl
19	74	50.0	15	12	US-10-261-446-19	Sequence 19, Appl
20	74	50.0	15	12	US-10-239-313A-618	Sequence 618, App
21	74	50.0	15	15	US-10-204-362-4	Sequence 7, Appl
22	74	50.0	15	15	US-10-223-711-7	Sequence 4, Appl
23	74	50.0	16	9	US-09-848-834A-2	Sequence 2, Appl
24	74	50.0	16	12	US-09-930-915A-64	Sequence 64, Appl
25	74	50.0	17	11	US-09-865-294-3	Sequence 3, Appl
26	74	50.0	17	12	US-10-239-313A-619	Sequence 619, App
27	74	50.0	19	12	US-10-239-313A-620	Sequence 620, App
28	74	50.0	29	9	US-09-732-754-1	Sequence 1, Appl
29	74	50.0	29	12	US-10-161-760-1	Sequence 1, Appl
30	74	50.0	29	15	US-10-206-535-1	Sequence 15, Appl
31	74	50.0	31	9	US-09-848-834A-15	Sequence 145, App
32	74	50.0	872	12	US-10-241-596-145	Sequence 145, App
33	74	50.0	879	12	US-10-241-596-143	Sequence 143, App
34	74	50.0	887	12	US-10-241-596-147	Sequence 147, App
35	74	50.0	1315	12	US-10-241-596-141	Sequence 141, App
36	73	49.3	29	15	US-10-238-607-33	Sequence 33, Appl
37	73	49.3	29	15	US-10-238-607-34	Sequence 34, Appl
38	73	49.3	31	9	US-09-848-834A-9	Sequence 9, Appl
39	73	49.3	34	9	US-09-848-834A-10	Sequence 10, Appl
40	73	49.3	47	9	US-09-848-834A-17	Sequence 17, Appl
41	73	49.3	50	9	US-09-848-834A-18	Sequence 18, Appl
42	72	48.6	50	9	US-09-943-548-9	Sequence 9, Appl
43	72	48.6	50	15	US-10-339-522-9	Sequence 9, Appl
44	70	47.3	14	12	US-09-932-165-1481	Sequence 1481, Ap
45	70	47.3	14	12	US-10-116-118-32	Sequence 32, Appl

## ALIGNMENTS

```
RESULT 1
US-09-848-834a-11
: Sequence 11, Application US/09848834A
: Patent No. US20020076416A1
: GENERAL INFORMATION:
: APPLICANT: Aption Corporation
: TITLE OF INVENTION: Chimeric Peptide Immunogens
: FILE REFERENCE: 1102865-0047
: CURRENT APPLICATION NUMBER: US/09/848,834A
: PRIOR FILING DATE: 2001-05-04
: PRIOR APPLICATION NUMBER: 60/202,328
: PRIOR FILING DATE: 2000-05-05
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 11
: LENGTH: 28
: TYPE: PRT
: FEATURE:
: ORGANISM: Artificial Sequence
: OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 830-844 of
: OTHER INFORMATION: Tetanus Toxoid Precursor (Tentoxylisin) linked by a spacer to
: OTHER INFORMATION: Ino acid sequence 2-10 of the GnRH hormone
: NAME/KEY: MOD_RES
: LOCATION: (1)..(1)
: OTHER INFORMATION: Amidated-glutamine
: NAME/KEY: MOD_RES
: LOCATION: (28)..(28)
: OTHER INFORMATION: Amidated-glycine or glycylamide
: NAME/KEY: PEPTIDE
: LOCATION: (1)..(15)
: OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus Toxoid Precursor
: OTHER INFORMATION: (Tentoxylisin)
: NAME/KEY: PEPTIDE
: LOCATION: (16)..(19)
: OTHER INFORMATION: Spacer peptide
: NAME/KEY: PEPTIDE
: LOCATION: (20)..(28)
: OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
```

US-09-848-834A-11

Query Match 99.3%; Score 147; DB 9; Length 28;  
Best Local Similarity 100.0%; Pred. No. 6,1e-15;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OYIKANSKFIGITELGPSLHWSYGLRP 27  
DB 1 OYIKANSKFIGITELGPSLHWSYGLRP 27

## RESULT 2

US-09-848-834A-19  
Sequence 19, Application US/09848834A  
Patent No. US20020076416A1  
GENERAL INFORMATION:  
APPLICANT: Aphlon Corporation  
TITLE OF INVENTION: Chimeric Peptide Immunogens  
FILE REFERENCE: 1102865-0047  
CURRENT APPLICATION NUMBER: US/09/848, 834A  
CURRENT FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: 60/202,328  
PRIOR FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 19

LENGTH: 46  
TYPE: PRT  
ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human  
OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetanus  
OTHER INFORMATION: toxoid precursor (Tetnoxylysin) linked by a spacer to amino acid  
NAME/KEY: MOD\_RES  
LOCATION: (1)..(46)  
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
NAME/KEY: MOD\_RES  
LOCATION: (46)..(46)  
OTHER INFORMATION: Amidated glycine or glycylamide  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(10)  
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
NAME/KEY: PEPTIDE  
LOCATION: (11)..(16)  
OTHER INFORMATION: Spacer peptide  
NAME/KEY: PEPTIDE  
LOCATION: (17)..(31)  
OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor  
NAME/KEY: PEPTIDE  
LOCATION: (32)..(37)  
OTHER INFORMATION: Spacer peptide  
NAME/KEY: PEPTIDE  
LOCATION: (38)..(46)  
OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-19

Query Match 91.9%; Score 136; DB 9; Length 46;  
Best Local Similarity 93.1%; Pred. No. 4,4e-13;  
Matches 27; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 OYIKANSKFIGITELG--GPSLHWSYGLRP 27  
DB 17 OYIKANSKFIGITELSGPSLHWSYGLRP 45

## RESULT 3

US-10-076-674-7  
Sequence 7, Application US/10076674  
Publication No. US20030165478A1  
GENERAL INFORMATION:  
APPLICANT: Sokoll, Kenneth K.

TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System  
FILE REFERENCE: Immunogen Delivery System  
CURRENT APPLICATION NUMBER: US/10/076,674  
CURRENT FILING DATE: 2002-04-23  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7

LENGTH: 27  
TYPE: PRT  
ORGANISM: Human  
US-10-076-674-7

Query Match 76.0%; Score 112.5; DB 12; Length 27;  
Best Local Similarity 85.2%; Pred. No. 7,8e-10;  
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 OYIKANSKFIGITELGPSLHWSYGLRP 27  
DB 3 OYIKANSKFIGITELG--HWSYGLRP 26

## RESULT 4

US-10-237-656-1  
Sequence 1, Application US/10237656  
Publication No. US2003015281A1  
GENERAL INFORMATION:  
APPLICANT: SAINT-REMY, Jean-Marie  
TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE TREATMENT OF  
FILE REFERENCE: 2002-0771/LC/01699  
CURRENT APPLICATION NUMBER: US/10/237,656  
CURRENT FILING DATE: 2002-09-10  
PRIOR APPLICATION NUMBER: 09/362,731  
PRIOR FILING DATE: 1999-07-29  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Version 2.1  
SEQ ID NO 1  
LENGTH: 31  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide  
US-10-237-656-1

Query Match 55.1%; Score 81.5; DB 12; Length 31;  
Best Local Similarity 61.3%; Pred. No. 3,6e-05;  
Matches 19; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

QY 1 OYIKANSKFIGITELG-----PSLHWS 22  
DB 1 OYIKANSKFIGITELGHEIKKVLVPCGHGS 31

## RESULT 5

US-09-848-834A-12  
Sequence 12, Application US/09848834A  
Patent No. US20020076416A1  
GENERAL INFORMATION:  
APPLICANT: Aphlon Corporation  
TITLE OF INVENTION: Chimeric Peptide Immunogens  
FILE REFERENCE: 1102865-0047  
CURRENT APPLICATION NUMBER: US/09/848, 834A  
CURRENT FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: 60/202,328  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 12  
LENGTH: 33  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 378-398 of

OTHER INFORMATION: Plasmodium falciparum CSP protein linked by a spacer to amino ac  
OTHER INFORMATION: id sequence 2-10 of the GnRH hormone  
NAME/KEY: MOD\_RES  
LOCATION: (1)..(1)  
OTHER INFORMATION: Amidated aspartic acid  
NAME/KEY: MOD\_RES  
LOCATION: (33)..(33)  
OTHER INFORMATION: Amidated glycine or glycylamide  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(20)  
OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria  
OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite  
NAME/KEY: PEPTIDE  
LOCATION: (21)..(24)  
OTHER INFORMATION: Spacer peptide  
NAME/KEY: PEPTIDE  
LOCATION: (25)..(33)  
OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-12

Query Match 55.1%; Score 81.5; DB 9; Length 33;  
Best Local Similarity 66.7%; Pred. No. 3.9e-05;  
Matches 16; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 4 KANSKFIGITELGPSLHMSYGLRP 27  
DB 10 KASVSF-NVNSGPSLHMSYGLRP 32

RESULT 6  
US-10-237-656-5  
Sequence 5; Application US/10237656  
Publication No. US20030152581A1  
GENERAL INFORMATION:  
APPLICANT: SAINT-REMY, Jean-Marie  
TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE TREATMENT OF AL  
FILE REFERENCE: 2002-0771/LC/01699  
CURRENT APPLICATION NUMBER: US/10/237,656  
PRIOR FILING DATE: 2002-09-10  
PRIOR APPLICATION NUMBER: 09/362,731  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Version 2.1  
SEQ ID NO 5  
LENGTH: 32  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide  
US-10-237-656-5

Query Match 54.1%; Score 80; DB 12; Length 32;  
Best Local Similarity 100.0%; Pred. No. 6.3e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16  
DB 1 QYIKANSKFIGITELG 16

RESULT 7  
US-10-237-656-3  
Sequence 3; Application US/10237656  
Publication No. US20030152581A1  
GENERAL INFORMATION:  
APPLICANT: SAINT-REMY, Jean-Marie  
APPLICANT: JACQUEMIN, Marc  
TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE TREATMENT OF AL  
FILE REFERENCE: 2002-0771/LC/01699  
CURRENT APPLICATION NUMBER: US/10/237,656  
CURRENT FILING DATE: 2002-09-10

PRIOR APPLICATION NUMBER: 09/362,731  
PRIOR FILING DATE: 1999-07-29  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Version 2.1  
SEQ ID NO 3  
LENGTH: 137  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide  
US-10-237-656-3

Query Match 54.1%; Score 80; DB 12; Length 137;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16  
DB 2 QYIKANSKFIGITELG 17

RESULT 8  
US-09-943-548-8  
Sequence 8; Application US/09943548  
Patent No. US20020042364A1  
GENERAL INFORMATION:  
APPLICANT: Rittershaus, Charles W.  
TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVI  
FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2  
CURRENT APPLICATION NUMBER: US/09/943,548  
PRIOR FILING DATE: 2001-08-30  
PRIOR APPLICATION NUMBER: 08/432,483  
PRIOR FILING DATE: 1995-05-01  
PRIOR APPLICATION NUMBER: PCT/US96/06147  
PRIOR FILING DATE: 1996-05-01  
PRIOR APPLICATION NUMBER: 08/945,289  
PRIOR FILING DATE: 1997-10-17  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Version 3.1  
SEQ ID NO 8  
LENGTH: 50  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: vaccine peptide of the invention  
US-09-943-548-8

Query Match 52.7%; Score 78; DB 9; Length 50;  
Best Local Similarity 94.1%; Pred. No. 0.0002; 1; Indels 0; Gaps 0;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGP 17  
DB 2 QYIKANSKFIGITELGP 18

RESULT 9  
US-10-339-522-8  
Sequence 8; Application US/10339522  
Publication No. US20030108559A1  
GENERAL INFORMATION:  
APPLICANT: Rittershaus, Charles W.  
APPLICANT: Thomas, Lawrence J.  
TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVI  
FILE REFERENCE: TCS-411.1P US-3  
CURRENT APPLICATION NUMBER: US/10/339,522  
PRIOR FILING DATE: 2003-01-08  
PRIOR APPLICATION NUMBER: 08/432,483  
PRIOR FILING DATE: 1995-05-01  
PRIOR APPLICATION NUMBER: PCT/US96/06147  
PRIOR FILING DATE: 1996-05-01  
PRIOR APPLICATION NUMBER: 08/945,289

PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 09/943,334  
PRIOR FILING DATE: 2001-08-30  
PRIOR APPLICATION NUMBER: 09/943,548  
PRIOR FILING DATE: 2001-08-30  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 8  
LENGTH: 50  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: vaccine peptide of the invention  
US-10-339-522-8

Query Match 52.7%; Score 78; DB 15; Length 50;  
Best Local Similarity 94.1%; Pred. No. 0.0002;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYKANSKFITGLTLP 17  
DB 2 OYKANSKFITGLTLP 18

RESULT 10  
US-10-128-711-110  
Sequence 110, Application US/10128711  
Publication No. US20030099634A1  
GENERAL INFORMATION:  
APPLICANT: VITIELLO, Maria A.  
CHESTNUT, Robert W.  
SETTE, Alessandro D.  
CELIS, Esteban  
GRAY, Howard  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
NUMBER OF SEQUENCES: 153  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourile and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/128,711  
FILING DATE: 22-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/197,484  
FILING DATE: 16-FEB-1994  
APPLICATION NUMBER: US 07/935,811  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992  
APPLICATION NUMBER: US 07/827,682  
FILING DATE: 29-JAN-1992  
APPLICATION NUMBER: US 07/749,568  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (206) 623-6793  
INFORMATION FOR SEQ ID NO: 110:  
SEQUENCE CHARACTERISTICS:

LENGTH: 24 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 110:  
US-10-128-711-110

Query Match 52.0%; Score 77; DB 15; Length 24;  
Best Local Similarity 88.9%; Pred. No. 0.00013;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 OYKANSKFITGLTLP 18  
DB 1 OYKANSKFITGLTLP 18

RESULT 11  
US-09-848-834A-20  
Sequence 20, Application US/09848834A  
Patent No. US20020076416A1  
GENERAL INFORMATION:  
APPLICANT: Apton Corporation  
TITLE OF INVENTION: Chimeric Peptide Immunogens  
FILE REFERENCE: 1102865-0047  
CURRENT APPLICATION NUMBER: US/09/848,834A  
CURRENT FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: 60/202,328  
PRIOR FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 20  
LENGTH: 51  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of hum  
OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plas  
NAME/KEY: MOD\_RES  
LOCATION: (1)..(1)  
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
NAME/KEY: MOD\_RES  
LOCATION: (51)..(51)  
OTHER INFORMATION: Amidated glycine or glycinamide  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(10)  
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
NAME/KEY: PEPTIDE  
LOCATION: (11)..(16)  
OTHER INFORMATION: Spacer peptide  
NAME/KEY: PEPTIDE  
LOCATION: (17)..(36)  
OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum  
NAME/KEY: PEPTIDE  
LOCATION: (37)..(42)  
OTHER INFORMATION: circumsporozoite (CSP) protein  
NAME/KEY: PEPTIDE  
LOCATION: (43)..(51)  
OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-20

Query Match 51.7%; Score 76.5; DB 9; Length 51;  
Best Local Similarity 64.0%; Pred. No. 0.00034;  
Matches 16; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

OY 4 KANSKF-IGITGLPSLHMSYGLRP 27  
DB 26 KASSVFNVNSSSGPSLHMSYGLRP 50

RESULT 12

```
US-09-943-548-2
; Sequence 2, Application US/09943548
; Patent No. US20020042364A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY
; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2
; CURRENT APPLICATION NUMBER: US/09/943,548
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vaccine peptide of the invention
US-09-943-548-2
```

```
Query Match          51.4%; Score 76; DB 9; Length 31;
Best Local Similarity 93.8%; Pred. No. 0.00024;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 OYIKANSKFIGITELG 16
        |||||
DB      2 OYIKANSKFIGITELG 17
```

```
RESULT 13
US-10-339-522-2
; Sequence 2, Application US/10339522
; Publication No. US20030108559A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY
; FILE REFERENCE: TCS-411.1P US-3
; CURRENT APPLICATION NUMBER: US/10/339,522
; PRIOR FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 09/943,334
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/943,548
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vaccine peptide of the invention
US-10-339-522-2
```

```
Query Match          51.4%; Score 76; DB 15; Length 31;
Best Local Similarity 93.8%; Pred. No. 0.00024;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 OYIKANSKFIGITELG 16
        |||||
DB      2 OYIKANSKFIGITELG 17
```

```
RESULT 14
US-09-862-849-2
; Sequence 2, Application US/09862849
; Patent No. US20020013274A1
; GENERAL INFORMATION:
; APPLICANT: Sudhir Paul
; APPLICANT: Larry J. Smith
; APPLICANT: Gennady Gololobov
; TITLE OF INVENTION: Methods for Identifying Inducers and Inhibitors of Proteolytic
; FILE REFERENCE: UNMC 63123 DIV
; CURRENT APPLICATION NUMBER: US/09/862,849
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: US 09/046,373
; PRIOR FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-862-849-2
```

```
Query Match          50.0%; Score 74; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 OYIKANSKFIGITEL 15
        |||||
DB      1 OYIKANSKFIGITEL 15
```

```
RESULT 15
US-09-785-215-4
; Sequence 4, Application US/09785215
; Publication No. US20020187157A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, Martin Roland et al.
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
; FILE REFERENCE: 3631-0107P
; CURRENT APPLICATION NUMBER: US/09/785,215
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-785-215-4
```

```
Query Match          50.0%; Score 74; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 OYIKANSKFIGITEL 15
        |||||
DB      1 OYIKANSKFIGITEL 15
```

Search completed: September 24, 2003, 17:33:07  
Job time : 14.0388 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:30:06 ; Search time 9.85042 Seconds  
(without alignments)  
273.361 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148  
Sequence: 1 QYKANSKEFIGITELGSLHWSGLRXPX 28

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	50.0	1315	1 BTCLTN	tentoxilysin (EC 3
2	54	36.5	67	2 I78541	gonadoliberin prec
3	54	36.5	90	1 RHMSG	gonadoliberin prec
4	54	36.5	92	1 RHMG	gonadoliberin prec
5	54	36.5	92	1 RHMG	gonadoliberin prec
6	52	35.1	10	1 RHMG	gonadoliberin - sh
7	52	35.1	10	1 RHMG	gonadoliberin - sh
8	52	35.1	89	2 I51423	gonadoliberin prec
9	51.5	34.8	824	2 AD3098	periplasmic nitrat
10	51.5	34.8	834	2 E98188	periplasmic nitrat
11	51	34.5	388	2 C72710	probable fmu prote
12	50.5	34.1	244	2 S29882	class II histocomp
13	50	33.8	274	2 T39087	hypothetical prote
14	50	33.8	480	2 A83487	hypothetical prote
15	49	33.1	452	2 T38962	probable MRS trans
16	49	33.1	484	2 E83245	hypothetical prote
17	49	33.1	3133	2 S52093	potassium uptake p
18	48	32.4	10	1 RHA01	hemocytin - silkw
19	48	32.4	10	1 RHA01	gonadoliberin I -
20	48	32.4	318	2 AD2810	gonadoliberin I pr
21	48	32.4	318	2 AD2810	transketolase limp
22	48	32.4	318	2 AD2810	probable transket
23	47.5	32.1	205	2 AB0294	pyruvate, water di
24	47.5	32.1	205	2 AB0294	conserved hypotet
25	46.5	31.4	193	2 F89967	hypothetical prote
26	46.5	31.4	245	2 S29880	class II histocomp
27	46.5	31.4	388	2 A82445	hypothetical prote
28	46	31.1	98	2 I50739	gonadotropin-relea
29	46	31.1	171	2 S38237	hypothetical prote

30	46	31.1	186	2 A90167	adenylate cyclase,
31	46	31.1	349	2 E75611	glucosamine-fructo
32	46	31.1	459	2 G82431	C4-dicarboxylate t
33	46	31.1	522	2 T44369	pyruvate, water di
34	46	31.1	780	2 D75361	phosphoenolpyruvat
35	46	31.1	4056	2 H96599	protein F14J16.10
36	45.5	30.7	322	2 T38399	probable amidohyd
37	45	30.4	256	2 T50616	hypothetical prote
38	45	30.4	258	2 F72052	peptidyl-prolyl cl
39	45	30.4	258	2 B86573	FKBP-type peptidyl
40	45	30.4	368	2 T27432	hypothetical prote
41	45	30.4	397	2 B70815	probable transamin
42	45	30.4	435	2 C89857	conserved hypotet
43	45	30.4	456	2 S55661	hypothetical prote
44	45	30.4	644	2 S46746	hypothetical prote
45	45	30.4	674	2 H72423	alpha-glucuronidas

#### ALIGNMENTS

##### RESULT 1

BTCLTN  
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani

N:Alternate names: tetanus neurotoxin

C:Species: Clostridium tetani

C>Date: 31-Mar-1988 #sequence, revision 31-Mar-1988 #text, change 03-Jun-2002

C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364

R:Eisel, U.; Jarasch, W.; Goretski, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, EMO J. 5, 2495-2502, 1986

A>Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b

A:Reference number: A25689; M0ID:87053814; PMID:3536478

A:Accession: A25689

A:Molecule type: DNA

A:Residues: 1-1315 <EIS>

A:Cross-references: GB:X04436; NID:q40769; PIDN:CAA2803.1; PID:q40770

R:Fairweather, N.F.; Lyness, V.A.

Nucleic Acids Res. 14, 7809-7812, 1986

A>Title: The complete nucleotide sequence of tetanus toxin.

A:Reference number: A25757; M0ID:87040747; PMID:3774547

A:Accession: A25757

A:Molecule type: DNA

A:Residues: 1-1315 <FAI>

A:Cross-references: GB:X06214; NID:q40773; PIDN:CAA29564.1; PID:q40774

R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.

J. Bacteriol. 165, 21-27, 1986

A>Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C 1

A:Reference number: A25194; M0ID:86085672; PMID:3510187

A:Accession: A25194

A:Molecule type: DNA

A:Residues: 743-1315 <FA2>

A:Cross-references: GB:M12739; NID:q144920; PIDN:AAA23282.1; PID:q144921

A:Accession: B25194

A:Molecule type: protein

A:Residues: 865-894 <FA3>

A:Reference number: A25194; M0ID:86085672; PMID:3510187

R:Matsumoto, M.; Ichi, D.L.; Sugimoto, N.; Ozutsuni, K.; Okabe, T.

Infect. Immun. 57, 3588-3593, 1989

A>Title: Isolation, purification, and characterization of fragment B, the NH-2-termin

A:Reference number: A60759; M0ID:90035436; PMID:2478476

A:Accession: A60759

A:Molecule type: protein

A:Residues: 461-475 <MAP>

R:Demetz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.

J. Immunol. 142, 394-402, 1989

A>Title: Delination of several DR-restricted tetanus toxin T cell epitopes.

A:Reference number: J50098; M0ID:89093918; PMID:2463305

A:Contents: annotation; epitope region

R:Schlavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta,

Nature 359, 832-835, 1992

A>Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo

A:Reference number: S27125; M0ID:93063293; PMID:1331807

A:Contents: annotation



C:Genetics:  
A:Gene: GDB:GNH: LHRH; GRH  
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760  
A:Map position: Bp21-8p11.2  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyrogutamyl acid  
F:1-23/Domin: signal sequence #status predicted <SIG>  
F:24-92/Product: progonaoliberin #status predicted <GNG>  
F:24-33/Product: gonadoliberin #status experimental <MAT>  
F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 36.5%; Score 54; DB 1; Length 92;  
Best Local Similarity 90.0%; Pred. No. 0.96;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27  
| | | | | | | | | |  
Db 23 SQHWSYGLRP 32

RESULT 5  
RHRTG  
gonadoliberin precursor - rat  
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormo  
N:Contains: gonadoliberin; prolactin release-inhibiting factor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 18-Jun-1999  
C:Accession: A40147; B26173; A48410  
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.  
Mol. Endocrinol. 3, 1257-1262, 1989  
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex  
A:Reference number: A40147; MUID:89384661; PMID:2476659  
A:Accession: A40147  
A:Residues: 1-92 <BON>  
A:Molecule type: DNA  
A:Cross-references: GB:M12579; NID:g204447; PIDN:AAA41264.1; PID:g204448  
R:Adelman, J.P.; Mason, A.U.; Hayflick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadob  
A:Reference number: A94090; MUID:86094338; PMID:2867548  
A:Accession: B26173  
A:Molecule type: mRNA  
A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12579; NID:g204445; PIDN:AAA41263.1; PID:g204446  
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Bialock, J.E.  
Cell. Mol. Neurobiol. 12, 447-454, 1992  
A:Title: Thyrocytes express a mRNA that is identical to hypothalamic luteinizing hormone  
A:Reference number: A48410; MUID:93105480; PMID:1468115  
A:Accession: A48410  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-92 <MAI>  
A:Cross-references: GB:550870; NID:g262059; PIDN:AAB24572.1; PID:g262060  
A:Experimental source: thymus  
A:Note: Sequence extracted from NCBI backbone (NCBIN:121082, NCBIP:121083)  
C:Genetics:  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: stimulates pituitary secretion of luteotropin and follitropin  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyrogutamyl acid; p  
F:1-23/Domin: signal sequence #status predicted <SIG>  
F:24-92/Product: progonaoliberin #status predicted <GNG>  
F:24-33/Product: gonadoliberin #status predicted <GNG>  
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted

F;33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match	36.5%	Score 54	DB 1	Length 92
Best Local Similarity	90.0%	Pred. No. 0.96		
Matches	9	Conservative	0	Mismatches 1
				Indels 0
				Gaps 0

QY	18	SLHWSYGLRP	27
Db	23	SQHWSYGLRP	32

RESULT 6  
RHPGG

```

C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1991 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Accession: A01411
A:Reference number: A90172; MUID:72114303; PMID:4946067

A:Molecule type: protein
A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase method.
A:Reference number: A90176; MUID:72065376; PMID:4942726

A:Contents: annotation; synthesis
A>Note: the synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544; PMID:4946275

A:Contents: annotation
A>Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1,Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.1%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
|||||||
Db 2 HWSYGLRP 9

RESULT 7
RHSHG
gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amos, M.; Ling, N.; Morahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314; PMID:4550508
A:Accession: A93780
A:Molecule type: protein
A:Residues: 1-10 <BBS>
A>Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1,Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

```

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27  
 |||||  
 Db 2 HWSYGLRP 9

# RESULT 8 151423

gonadoliberin precursor - African clawed frog  
 N:Alternate names: luteinizing hormone releasing hormone  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
 C:Accession: 151423  
 R:Hayes, W.P.; Wray, S.; Battey, J.F.  
 Endocrinology 134, 1835-1845, 1994  
 A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved domain  
 A:Reference number: 151423; PMID:94185563; PMID:8137750  
 A:Accession: 151423  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-89 <HAY>  
 A:Cross-references: GB:I28040; NID:g496291; PIDN:AAA49728.1; PID:g496292  
 C:Genetics:  
 A:Gene: GnRH-I  
 C:Superfamily: gonadoliberin

Query Match 35.1%; Score 52; DB 2; Length 89;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27  
 |||||  
 Db 25 HWSYGLRP 32

# RESULT 9

AD3098  
 periplasmic nitrate reductase large subunit [imported] - Agrobacterium tumefaciens (strain C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C:Species: Agrobacterium tumefaciens  
 C:Accession: AD3098  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, D.; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; Mclellan, Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:21608550; PMID:11743193  
 A:Accession: AD3098  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-824 <KUR>  
 A:Cross-references: GB:AE008689; PIDN:AAL45202.1; PID:g17742881; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: napA  
 A:Map position: linear chromosome  
 C:Superfamily: Alcaligenes eutrophus nitrate reductase A chain

Query Match 34.8%; Score 51.5; DB 2; Length 824;  
 Best Local Similarity 37.0%; Pred. No. 23;  
 Matches 10; Conservative 9; Mismatches 1; Indels 7; Gaps 2;

QY 2 YIKANSKFI-GITELGPSLHWSYGLRP 27  
 ::::: |||  
 Db 279 FVRNHTKFRVGVTDIG-----YGLRP 299

RESULT 10  
 E98188  
 periplasmic nitrate reductase precursor napA (AF040988) [imported] - Agrobacterium tumefaciens

C:Species: Agrobacterium tumefaciens  
 C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 18-Nov-2002  
 C:Accession: E98188  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium  
 A:Reference number: A97359; PMID:21608551; PMID:11743194  
 A:Accession: E98188  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-834 <KUR>  
 A:Cross-references: GB:AE007870; PIDN:AAK89031.1; PID:g15158825; GSPDB:GN00170  
 C:Genetics:  
 A:Gene: AGR\_L\_917  
 A:Map position: linear chromosome  
 C:Superfamily: Alcaligenes eutrophus nitrate reductase A chain

Query Match 34.8%; Score 51.5; DB 2; Length 834;  
 Best Local Similarity 37.0%; Pred. No. 24;  
 Matches 10; Conservative 9; Mismatches 1; Indels 7; Gaps 2;

QY 2 YIKANSKFI-GITELGPSLHWSYGLRP 27  
 ::::: |||  
 Db 289 FVRNHTKFRVGVTDIG-----YGLRP 309

# RESULT 11

C72710  
 Probable fru protein APE1098 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
 C:Accession: C72710  
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Tawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J. DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero  
 A:Reference number: A72450; PMID:99310339; PMID:10382966  
 A:Accession: C72710  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-388 <KAW>  
 A:Cross-references: DBJ:AP000060; NID:g5104188; PIDN:BAA80083.1; PID:d1043869; PID:g  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE1098

Query Match 34.5%; Score 51; DB 2; Length 388;  
 Best Local Similarity 71.4%; Pred. No. 12;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 14 ELGPSLHWSYGLRP 27  
 ||| | |||  
 Db 345 ELGFRLLTWSYGLRP 358

# RESULT 12

S29982  
 class II histocompatibility antigen - Atlantic salmon  
 C:Species: Salmo salar (Atlantic salmon)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
 C:Accession: S29982  
 R:Hordvik, I.  
 submitted to the EMBL Data Library, October 1992  
 A:Reference number: S29980  
 A:Accession: S29982  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-244 <HOR>  
 A:Cross-references: EMBL:X70166; NID:g64369; PID:g64370  
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 34.1%; Score 50.5; DB 2; Length 244;

Best Local Similarity 57.9%; Pred. No. 8.8;  
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

OY 1 OYIKANS---KFTGTELG 16  
:||:| | |||  
Db 51 EXTRFNSTVGKFGVTELG 69

## RESULT 13

T39087  
hypothetical protein SPAC7D4.09c - fission yeast (*Schizosaccharomyces pombe*)

C/Species: *Schizosaccharomyces pombe*

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C/Accession: T39087

R/Genies: S.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, September 1997

A/Reference number: Z21826

A/Accession: T39087

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-274 <GEN>

A/Cross-references: EMBL:Z99532; PIDN:CAB16726.2; GSPDB:GN00066; SPDB:SPAC7D4.09c

A/Experimental source: strain 972h-; cosmid c7D4

C/Genetics:

A/Map position: 1

Query Match 33.8%; Score 50; DB 2; Length 274;  
Best Local Similarity 29.4%; Pred. No. 12;  
Matches 10; Conservative 7; Mismatches 9; Indels 8; Gaps 1;

OY 2 YIKANSKFTGTELGPSLHW-----SYGLRP 27  
| : | : | | | |  
Db 18 YFTSTLFVSLKNAPSLMKYGHDFGLKP 51

## RESULT 14

A83487  
probable MFS transporter PA1262 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C/Species: *Pseudomonas aeruginosa*

C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C/Accession: A83487

R/Stover, C.K.; Pham, X.Q.; Ertwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Llm,

Loxy, S.; Olson, M.V.

Nature 406, 959-964, 2000

A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho

A/Reference number: A82950; M0ID:20437337; PMID:10984043

A/Accession: A83487

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-480 <STO>

A/Cross-references: GB:AE004556; GB:AE004091; NID:g9947194; PIDN:AA04651.1; GSPDB:GN001

A/Experimental source: strain PA01

C/Genetics:

A/Map position: 1

Query Match 33.8%; Score 50; DB 2; Length 480;  
Best Local Similarity 88.9%; Pred. No. 22;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 15 LGPSLHWSY 23  
|||||||  
Db 224 LGPSLHWSW 232

## RESULT 15

T38962

hypothetical protein SPAC5D6.04 - fission yeast (*Schizosaccharomyces pombe*)

C/Species: *Schizosaccharomyces pombe*

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C/Accession: T38962

R/Skelton, J.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1997

A/Reference number: Z21820

A/Accession: T38962

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-452 <SKE>

A/Cross-references: EMBL:Z98056; PIDN:CAB10852.1; GSPDB:GN00066; SPDB:SPAC5D6.04

A/Experimental source: strain 972h-; cosmid c5D6

C/Genetics:

A/Map position: 1

A/Introns: 48/3; 92/2; 122/3

Query Match 33.1%; Score 49; DB 2; Length 452;  
Best Local Similarity 57.1%; Pred. No. 28;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 13 TELGPSLHWSYGLR 26  
:||:| | |||  
Db 166 SGLGALRWSYGLR 179

Search completed: September 24, 2003, 17:49:31  
Job time : 10.8504 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:00:31 ; Search time 5.19668 Seconds  
(without alignments)  
253.383 Million cell updates/sec

Title: US-09-848-834A-11  
Perfect score: 148  
Sequence: 1 QYIKANSKFITGLTGLPSLHWSYGLRPX 28

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	74	50.0	1314	1	TEXX_CLOTE
2	54	36.5	67	1	GONI_MACMO
3	54	36.5	90	1	GONI_MOUSE
4	54	36.5	91	1	GONI_PIG
5	54	36.5	92	1	GONI_HUMAN
6	54	36.5	92	1	GONI_RAT
7	52	35.1	61	1	GONI_SHEEP
8	52	35.1	63	1	GONI_MESAU
9	52	35.1	89	1	GONI_XENLA
10	52	35.1	92	1	GONI_TUPGB
11	49	33.1	452	1	YD04_SCHPO
12	49	33.1	3133	1	HMCT_BOMMO
13	48	32.4	10	1	GONI_MLMT
14	48	32.4	92	1	GONI_CAVPO
15	48	32.4	92	1	GONI_CHICK
16	47	31.8	216	1	YD04_LEPIN
17	46	31.1	94	1	GONI_HAPBU
18	46	31.1	95	1	GONI_MORSA
19	46	31.1	95	1	GONI_PAGMA
20	46	31.1	95	1	GONI_SPAUV
21	46	31.1	99	1	GONI_DICLA
22	46	31.1	110	1	YHBJ_ACTAC
23	46	31.1	459	1	DCVC_VIRCH
24	46	31.1	780	1	PRSA_DEIRA
25	45.5	30.7	322	1	YAOB_SCHPO
26	45	30.4	258	1	MIP_CHUPN
27	45	30.4	347	1	ALIC_MOUSE
28	45	30.4	644	1	YHJ9_YEAST
29	45	30.4	674	1	AGUA_THEMEA
30	45	30.4	831	1	NAPA_ALCEU
31	44	29.7	66	1	VG84_BPMLS
32	44	29.7	90	1	GON8_RANDY
33	44	29.7	255	1	YABD_BACSU

34	44	29.7	256	1	YD83_METUA	Q58778 methanococ
35	44	29.7	293	1	CFXO_GUTH	078450 guillardi
36	44	29.7	357	1	YOC2_CAEEL	009305 caenorhabdi
37	44	29.7	459	1	YGM9_YEAST	P53083 saccharomyc
38	44	29.7	613	1	FIBP_ADEMI	P15721 mouse adeno
39	44	29.7	664	1	ACES_ANOST	P56161 anopheles s
40	44	29.7	1016	1	PDQG_ECOLI	P32176 escherichia
41	43.5	29.4	90	1	GON3_DICLA	091a09 dicentrarch
42	43.5	29.4	231	1	CYSH_BACHD	09cct3 bacillus ha
43	43.5	29.4	236	1	CYH2_BACSU	006737 bacillus su
44	43	29.1	185	1	PTH_RICPR	09c6v4 rickettsia
45	43	29.1	193	1	CIT1_HUMAN	099966 homo sapien

## ALIGNMENTS

RESULT 1	ID	TEXX_CLOTE	STANDARD:	PRT:	1314 AA.
AC	P04958:				
DT	13-AUG-1987 (Rel. 05, Last sequence update)				
DT	13-AUG-1987 (Rel. 05, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains:				
DE	Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy				
DE	chain (Tetanus toxin chain H)].				
GN	TEXX OR CTP60.				
OS	Clostridium tetani.				
OG	Plasmid pE88, and Plasmid 75 kbp.				
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;				
OX	Clostridium.				
OX	NCBI_TaxID=1513;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	PLASMID=75 Kbp;				
RX	MEDLINE=87053814; PubMed=3536478;				
RA	Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,				
RA	Weller U., Hudek M., Habermann E., Niemann H.;				
RT	"Tetanus toxin: primary structure, expression in E. coli, and				
RT	homology with botulinum toxins.";				
EMBO	J. 5:2495-2502(1986).				
[2]					
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CN3911; PLASMID=75 Kbp;				
RX	MEDLINE=87040747; PubMed=3774547;				
RA	Fairweather N.F., Lyness V.A.;				
RT	"The complete nucleotide sequence of tetanus toxin.";				
RL	Nucleic Acids Res. 14:7809-7812(1986).				
[3]					
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Massachusetts / E88; PLASMID=pE88;				
RX	MEDLINE=22457253; PubMed=12552129;				
RA	Brueggemann H., Baeumer S., Fricke W.F., Wlezer A., Ilesgang H.,				
RA	Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,				
RA	Gottschalk G.;				
RT	"The genome sequence of Clostridium tetani, the causative agent of				
RT	tetanus disease.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).				
[4]					
RP	SEQUENCE OF 742-1314 FROM N.A.				
RC	PLASMID=75 Kbp;				
RX	MEDLINE=86085672; PubMed=3510187;				
RA	Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;				
RT	"Cloning, nucleotide sequencing, and expression of tetanus toxin				
RT	fragment C in Escherichia coli.";				
RL	J. Bacteriol. 165:21-27(1986).				
[5]					
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.				
RX	MEDLINE=90201034; PubMed=2108021;				
RA	Kriegstein K., Henschen A., Weller U., Habermann E.;				
RT	"Arrangement of disulfide bridges and positions of sulfhydryl groups				
RT	in tetanus toxin.";				

RL Eur. J. Biochem. 188:39-45(1990).  
 RN [6]  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE=92037649; PubMed=1935979;  
 RA Krugstein K.G., Henschen A.H., Weller U., Habermann E.;  
 RT "Limited proteolysis of tetanus toxin. Relation to activity and  
 RL identification of cleavage sites.";  
 RL Eur. J. Biochem. 202:41-51(1991).  
 RN [7]  
 RP IDENTIFICATION AS ZINC-PROTEASE.  
 RX MEDLINE=93010948; PubMed=1396558;  
 RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,  
 RA Montecucco C.;  
 RT "Tetanus toxin is a zinc protein and its inhibition of  
 RT neurotransmitter release and protease activity depend on zinc";  
 RL EMBO J. 11:3577-3583(1992).  
 RN [8]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=93063293; PubMed=1331807;  
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,  
 RA Dasgupta B.R., Montecucco C.;  
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release  
 RT by proteolytic cleavage of synaptobrevin.";  
 RL Nature 353:832-835(1992).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.  
 RX MEDLINE=97475217; PubMed=9334741;  
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,  
 RA Sax M.;  
 RT "Structure of the receptor binding fragment HC of tetanus  
 RT neurotoxin.";  
 RL Nat. Struct. Biol. 4:788-792(1997).  
 CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-1-PHE-77  
 CC BOND OF SYNAPTOSOMAL-2.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-1-Phe-77 bond in  
 CC synaptobrevin 2.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (by similarity).  
 CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO  
 CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE  
 CC AND ARE NON-TOXIC AFTER SEPARATION.  
 CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO  
 CC GANGLIOSIDE RECEPTORS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X04436; CAA28033.1; -;  
 DR EMBL: X06214; CAA29564.1; -;  
 DR EMBL: AF528097; AAO37454.1; -;  
 DR EMBL: M12739; AAA23282.1; -;  
 DR PIR: A25689; BTCLTN.  
 DR PDB: 1AF9; 29-APR-98.  
 DR PDB: 1A8D; 14-OCT-98.  
 DR PDB: 1D0H; 27-MAR-00.  
 DR PDB: 1DFO; 24-MAR-00.  
 DR PDB: 1DIW; 24-MAR-00.  
 DR PDB: 1DLJ; 24-MAR-00.  
 DR PDB: 1FV3; 05-SEP-01.  
 DR MEROPS: M27.001; -;  
 DR InterPro: IPR000395; Bontoxilysin.  
 DR InterPro: IPR006025; Zn\_MTEPdisse.  
 DR Pfam: PF01742; Peptidase\_M27; 1.

DR PRINTS: PR00760; BONTOTOXILYSIN.  
 DR PRODom: PD001963; Bontoxilysin; 1.  
 DR PROSITE: PS001442; ZINC\_PROTEASE; 1.  
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;  
 KW 3D-structure; Complete proteome.  
 FT INIT\_MET 0 0  
 FT CHAIN 1 456  
 FT CHAIN 1 456  
 FT METAL 232 233  
 FT ACT\_SITE 233 233  
 FT METAL 236 236  
 FT TRANSMEM 226 246  
 FT TRANSMEM 669 689  
 FT DISULFID 438 466  
 FT DISULFID 1076 1092  
 FT HELIX 876 882  
 FT TURN 883 883  
 FT TURN 884 891  
 FT STRAND 884 891  
 FT TURN 892 893  
 FT STRAND 894 897  
 FT STRAND 904 907  
 FT TURN 909 910  
 FT STRAND 912 915  
 FT STRAND 920 925  
 FT TURN 928 929  
 FT STRAND 932 935  
 FT STRAND 938 940  
 FT HELIX 941 946  
 FT TURN 949 956  
 FT STRAND 962 968  
 FT HELIX 969 970  
 FT TURN 972 977  
 FT STRAND 980 981  
 FT STRAND 983 985  
 FT HELIX 987 995  
 FT TURN 996 997  
 FT STRAND 998 1004  
 FT TURN 1006 1007  
 FT STRAND 1010 1016  
 FT STRAND 1020 1020  
 FT TURN 1021 1022  
 FT STRAND 1031 1037  
 FT TURN 1039 1040  
 FT STRAND 1042 1047  
 FT TURN 1048 1049  
 FT STRAND 1050 1056  
 FT TURN 1058 1059  
 FT STRAND 1068 1074  
 FT TURN 1079 1080  
 FT STRAND 1082 1091  
 FT HELIX 1097 1105  
 FT TURN 1106 1107  
 FT STRAND 1112 1112  
 FT STRAND 1114 1114  
 FT TURN 1116 1117  
 FT STRAND 1120 1120  
 FT STRAND 1122 1122  
 FT TURN 1123 1124  
 FT STRAND 1127 1131  
 FT HELIX 1132 1134  
 FT TURN 1135 1136  
 FT STRAND 1137 1141  
 FT TURN 1144 1145  
 FT STRAND 1148 1152  
 FT STRAND 1155 1158  
 FT TURN 1159 1162  
 FT STRAND 1163 1166  
 FT STRAND 1173 1178  
 FT TURN 1184 1185  
 FT STRAND 1188 1188  
 FT STRAND 1190 1190  
 FT TURN 1191 1192  
 FT STRAND 1193 1201  
 FT STRAND 1193 1201

Query Match 50.0%; Score 74; DB 1; Length 1314;  
 Best Local Similarity 100.0%; Pred. No. 0.004;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYKANSKEFIGTEL 15  
 DB 829 OYKANSKEFIGTEL 843

## RESULT 2

GONI\_MACMU STANDARD; PRT; 67 AA.  
 ID GONI\_MACMU  
 AC P55247;  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonadolibertin I precursor [contains: Gonadolibertin I (LH-RH I)  
 DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GNRH I) (Luliberin I); GNRH-associated, peptide I]  
 DE (Fragment).  
 GN GNRH1 OR GNRH OR LHRH.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecidae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hypothalamus;  
 RX MEDLINE=95124501; PubMed=7545971;  
 RA Ma Y.J., Costa M.E., Ojeda S.R.;  
 RT "Developmental expression of the genes encoding transforming growth  
 RT factor alpha and its receptor in the hypothalamus of female rhesus  
 RT macaques."  
 RL Neuroendocrinology 60:346-359(1994).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the GNRH family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

DR EMBL: S75918; AAB33096.1; -  
 DR PIR: I78541; I78541.  
 DR InterPro: IPR002012; GNRH.  
 DR InterPro: IPR004079; GonadolibertinI.  
 DR Pfam: PF00446; GNRH; 1.  
 DR PRINTS: PRO1541; GONADOLIBRNI.  
 DR PROSITE: PS00473; GNRH; 1.  
 KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal; Pyroglutamate carboxylic acid.  
 FT SIGNAL 1 1  
 FT NON\_TER 1 1  
 FT CHAIN <1 5 BY SIMILARITY.  
 FT PEPTIDE 6 15 PROGNADOLIBERTIN I.  
 FT PEPTIDE 19 15 GONADOLIBERTIN I.  
 FT ACT\_SITE 8 8 GNRH-ASSOCIATED PEPTIDE I.  
 FT MOD\_RES 6 6 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT MOD\_RES 6 6 ACTIVITY (BY SIMILARITY).  
 FT MOD\_RES 6 6 PYROGLUTAMATE CARBOXYLIC ACID (BY  
 FT MOD\_RES 15 15 SIMILARITY).  
 FT MOD\_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY  
 FT NON\_TER 67 67 SIMILARITY).  
 FT SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 36.5%; Score 54; DB 1; Length 67;  
 Best Local Similarity 90.0%; Pred. No. 0.21;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 18 SLHWSYGLRP 27  
 DB 5 SLHWSYGLRP 14

## RESULT 3

GONI\_MOUSE STANDARD; PRT; 90 AA.  
 ID GONI\_MOUSE  
 AC P13562;  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonadolibertin I precursor [contains: Gonadolibertin I (LH-RH I)  
 DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor  
 DE I].  
 GN GNRH1 OR GNRH.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87069928; PubMed=3024317;  
 RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,  
 RA Phillips H.S., Nikolic S., Seeburg P.H.;  
 RT "A deletion truncating the gonadotropin-releasing hormone gene is  
 RT responsible for hypogonadism in the hpg mouse."  
 RL Science 234:1366-1371(1986).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the GNRH family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

DR EMBL: M14872; AAA37717.1; -  
 DR PIR: A47578; RHMSG.  
 DR MGD: MGI:95789; Gnrh.  
 DR InterPro: IPR002012; GNRH.  
 DR InterPro: IPR004079; GonadolibertinI.  
 DR Pfam: PF00446; GNRH; 1.  
 DR PRINTS: PRO1541; GONADOLIBRNI.  
 DR PROSITE: PS00473; GNRH; 1.  
 KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Signal; Pyroglutamate carboxylic acid.  
 FT SIGNAL 1 21  
 FT CHAIN 22 90  
 FT PEPTIDE 22 90 PROGNADOLIBERTIN I.  
 FT PEPTIDE 31 90 GONADOLIBERTIN I.  
 FT ACT\_SITE 24 24 PROLACTIN RELEASE-INHIBITING FACTOR I.  
 FT MOD\_RES 22 22 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT MOD\_RES 31 31 ACTIVITY.  
 FT MOD\_RES 31 31 PYROGLUTAMATE CARBOXYLIC ACID.  
 FT MOD\_RES 90 90 AMIDATION (G-32 PROVIDE AMIDE GROUP).  
 FT SEQUENCE 90 AA; 10337 MW; IC0766FA4826AD9 CRC64;

Query Match 36.5%; Score 54; DB 1; Length 90;  
 Best Local Similarity 90.0%; Pred. No. 0.28;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 18 SLHWSYGLRP 27  
 1 | | | | | | | | | |

Db 21 SQHWSYGLRP 30

RESULT 4

ID	GONL_PIG	STANDARD:	PRT:	91 AA.
AC	P49921;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Progonadoliblerin I precursor [contains: Gonadoliblerin I (LH-RH I) (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].			
DE	GnRH OR GnRH.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Hypothalamus;			
RA	Weesner G.D., Matterl R.L., Becker B.A.;			
RT	Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE OF 24-33.			
RX	MEDLINE=72114303; PubMed=4946067;			
RA	Baba Y., Matsuo H., Schally A.V.;			
RT	"Structure of the porcine LH- and FSH-releasing hormone. II.			
RT	Confirmation of the proposed structure by conventional sequential analyses."			
RT	Biochem. Biophys. Res. Commun. 44:459-463(1971).			
RN	[3]			
RP	SYNTHESIS OF GONADOLIBERIN.			
RX	MEDLINE=72065376; PubMed=4942726;			
RA	Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;			
RT	"Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase method."			
RT	Biochem. Biophys. Res. Commun. 45:822-827(1971).			
RN	[4]			
RP	SYNTHESIS OF GONADOLIBERIN.			
RX	MEDLINE=72117544; PubMed=4946275;			
RA	Baba Y., Arimura A., Schally A.V.;			
RT	"On the tryptophan residue in porcine LH and FSH-releasing hormone."			
RT	Biochem. Biophys. Res. Commun. 45:483-487(1971)			
RL	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- SIMILARITY: Belongs to the GnRH family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage is for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: L32864; AAA31066.1; -			
DR	InterPro: IPR002012; GnRH.			
DR	InterPro: IPR004079; GonadoliblerinI.			
DR	Pfam: PF00446; GnRH; 1.			
DR	PRINTS: PR01541; GONADOLIBRNI.			
DR	PROSITE: PS00473; GnRH; 1.			
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Signal; Pyrrolidone carboxylic acid.			
FT	SIGNAL	1	23	
FT	CHAIN	24	91	PROGONADOLIBERIN I.
FT	PEPTIDE	24	33	GONADOLIBERIN I.
FT	ACT_SITE	26	26	APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
FT	MOD_RES	24	24	PYRROLIDONE CARBOXYLIC ACID.

FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).

SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;

Query Match

Best Local Similarity 90.0%; Score 54; DB 1; Length 91;

Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 18 SLHWSYGLRP 27

Db 23 SQHWSYGLRP 32

RESULT 5

ID	GONL_HUMAN	STANDARD:	PRT:	92 AA.
AC	P01148;			
DT	21-JUL-1966 (Rel. 01, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Progonadoliblerin I precursor [contains: Gonadoliblerin I (LH-RH I) (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated peptide I].			
DE	GnRH OR GnRH OR LHRH.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=8936682; PubMed=2671939;			
RA	Hayflick J.S., Adelman J.P., Seeburg P.H.;			
RT	"The complete nucleotide sequence of the human gonadotropin-releasing hormone gene."			
RT	Nucleic Acids Res. 17:6403-6403(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86094338; PubMed=2867548;			
RA	Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;			
RT	"Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat."			
RT	Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANT SER-16.			
RX	MEDLINE=85012739; PubMed=6090951;			
RA	Seeburg P.H., Adelman J.P.;			
RT	"Characterization of cDNA for precursor of human luteinizing hormone releasing hormone."			
RT	Nature 311:666-668(1984).			
RN	[4]			
RP	SEQUENCE OF 24-33.			
RX	MEDLINE=83126573; PubMed=6760865;			
RA	Tan L., Rousseau P.;			
RT	"The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in the human placenta."			
RT	Biochem. Biophys. Res. Commun. 109:1061-1071(1982).			
RN	[5]			
RP	VARIANT SER-16.			
RX	MEDLINE=99318093; PubMed=10391209;			
RA	Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N., Shaw N., Lane C.R., Lim E.P., Kalyanarman N., Nemesh J., Ziaugra L., Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O., Lander E.S.;			
RT	"Characterization of single-nucleotide polymorphisms in coding regions of human genes."			
RT	Nat. Genet. 22:231-238(1999).			
RL	[6]			
RP	ERRATUM.			
RA	Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N., Shaw N., Lane C.R., Lim E.P., Kalyanarman N., Nemesh J., Ziaugra L., Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O., Lander E.S.;			

```

RL  Nat. Genet. 23:373-373(1999).
CC  -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC  THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC  HORMONES.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC  Lutrelpse or Lutrelief (Ferring Pharmaceuticals) and Relisorm
CC  (Serono).
CC  -1- SIMILARITY: Belongs to the GnRH family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC  or send an email to license@isb-sdb.ch).
CC  -----
CC  EMBL; X01059; CAA25526.1; -
CC  EMBL; M12578; AAA35916.1; -
CC  EMBL; X15215; CAA33285.1; -
CC  PIR; S05308; RHHC.
CC  Genew; HGNC:4419; GNRH1.
CC  MIM; 152760; -
CC  DR  GO: 0005625; C:soluble fraction; TAS.
CC  DR  GO: 0005183; F:luteinizing hormone-releasing factor activity; TAS.
CC  DR  GO: 0007267; P:cell-cell signaling; TAS.
CC  DR  GO: 0007275; P:development; TAS.
CC  DR  GO: 0008285; P:negative regulation of cell proliferation; TAS.
CC  DR  GO: 0007165; P:signal transduction; TAS.
CC  DR  InterPro: IPR002012; GnRH.
CC  DR  InterPro: IPR004079; GonadolibertinI.
CC  DR  Pfam: PF00446; GnRH; 1.
CC  DR  PRINTS; PR01541; GONADOLIBERTIN.
CC  DR  PROSITE; PS00473; GNRH; 1.
CC  KW  Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC  Placenta; Pharmaceutical; Signal; Polymorphism;
CC  KW  Pyroglutamate carboxylic acid.
CC  FT  SIGNAL 1 23
CC  FT  CHAIN 24 92
CC  FT  PEPTIDE 24 33
CC  FT  ACT_SITE 26 26
CC  FT  ACT_SITE 26 26
CC  FT  MOD_RES 24 24
CC  FT  MOD_RES 33 33
CC  FT  VARIANT 16 16
CC  FT  SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;
CC  Query Match 36.5%; Score 54; DB 1; Length 92;
CC  Best Local Similarity 90.0%; Pred. No. 0.29;
CC  Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX  NCBI_TaxID=10116;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=86094338; PubMed=2867548;
RA  Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT  Isolation of the gene and hypothalamic cDNA for the common precursor
RT  of gonadotropin-releasing hormone and prolactin release-inhibiting
RT  factor in human and rat.;
RL  Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=89384661; PubMed=2476669;
RA  Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT  "The rat gonadotropin-releasing hormone: SH locus: structure and
RT  hypothalamic expression.";
RL  Mol. Endocrinol. 3:1257-1262(1989).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93105480; PubMed=1468115;
RA  Walter C.C., Marchetti B., Lebouef R.D., Blajock J.E.;
RT  "Thymocytes express a mRNA that is identical to hypothalamic
RT  luteinizing hormone-releasing hormone mRNA.";
RL  Cell. Mol. Neurobiol. 12:447-454(1992).
RN  [4]
RP  SEQUENCE OF 1-47 FROM N.A.
RC  TISSUE=Heart;
RX  MEDLINE=87149087; PubMed=3547652;
RA  Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT  "Two mammalian genes transcribed from opposite strands of the same
RT  DNA locus.";
RL  Science 235:1514-1517(1987).
CC  -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC  THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC  HORMONES.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC  -1- SIMILARITY: Belongs to the GnRH family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC  or send an email to license@isb-sdb.ch).
CC  -----
CC  EMBL; S50870; AAB24572.1; -
CC  DR  EMBL; M12579; AAA41263.1; -
CC  DR  EMBL; M31670; AAA41264.1; -
CC  DR  EMBL; M15527; AAA42141.1; ALT_SEQ.
CC  DR  EMBL; M15529; AAA42139.1; -
CC  DR  EMBL; M15528; -; NOT_ANNOTATED_CDS.
CC  DR  PIR; A40147; RHRTG.
CC  DR  InterPro: IPR002012; GnRH.
CC  DR  InterPro: IPR004079; GonadolibertinI.
CC  DR  Pfam: PF00446; GnRH; 1.
CC  DR  PRINTS; PR01541; GONADOLIBERTIN.
CC  DR  PROSITE; PS00473; GNRH; 1.
CC  KW  Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC  Placenta; Signal; Pyroglutamate carboxylic acid.
CC  FT  SIGNAL 1 23
CC  FT  CHAIN 24 92
CC  FT  PEPTIDE 24 33
CC  FT  ACT_SITE 26 26
CC  FT  ACT_SITE 26 26
CC  FT  MOD_RES 24 24
CC  FT  MOD_RES 33 33
CC  FT  SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;
CC  Query Match 36.5%; Score 54; DB 1; Length 92;

```

Best Local Similarity 90.0%; Pred. No. 0.29;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 18 SLHMSGGLRP 27  
| | | | | | | |  
DB 23 SOHMSGGLRP 32

RESULT 7  
GONI\_SHEEP STANDARD: PRT: 61 AA.

ID GONI\_SHEEP STANDARD: PRT: 61 AA.  
AC 028588;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Progonaoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]  
DE (Fragment).  
GN GNRI OR GNRI OR LHRI.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE OF 12-61 FROM N.A.  
RC STRAIN=western range; TISSUE=Hypothalamus;  
RA Rodriguez R.E., Wise M.E.;  
RL Submitted (OCT-1993) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 1-10.  
RX MEDLINE=72094314; PubMed=4550508;  
RA Burgess R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,  
RA Fellows R., Blackwell R., Vale W., Guillemin R.;  
RT "Primary structure of the ovine hypothalamic luteinizing hormone-  
RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass  
RT spectrometry-decapeptide-Edman degradation).";  
RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: Belongs to the GnRH family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC DR EMBL: U02517; AAA03433.1; -  
CC DR InterPro: IPR002012; GnRH.  
CC DR InterPro: IPR004079; GonadoliberinI.  
CC DR Pfam: PF00446; GnRH. 1.  
CC DR PRINTS: PR01541; GONADOLIBRNI.  
CC DR PROSITE: PS00473; GNRI. 1.  
CC KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
CC Placenta; Pyroglutamate carboxylic acid.  
CC FT CHAIN 1 1  
CC FT PEPTIDE 1 10  
CC FT PEPTIDE 14 >61  
CC FT ACT\_SITE 3 3  
CC FT MOD\_RES 1 1  
CC FT MOD\_RES 10 10  
CC FT NON\_TER 61  
CC SO SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;

Query Match 35.1%; Score 52; DB 1; Length 61;

Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 HMSGGLRP 27  
| | | | | | | |  
DB 2 HMSGGLRP 9

RESULT 8  
GONI\_MESAU STANDARD: PRT: 63 AA.

ID GONI\_MESAU STANDARD: PRT: 63 AA.  
AC 009163;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Progonaoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]  
DE (Fragment).  
GN GNRI OR GNRI OR LHRI.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Mesocricetus.  
OX NCBI\_TaxID=10036;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;  
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: Belongs to the GnRH family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC DR EMBL: U91938; AAB51302.1; -  
CC DR InterPro: IPR002012; GnRH.  
CC DR InterPro: IPR004079; GonadoliberinI.  
CC DR Pfam: PF00446; GnRH. 1.  
CC DR PRINTS: PR01541; GONADOLIBRNI.  
CC DR PROSITE: PS00473; GNRI. 1.  
CC KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
CC Placenta; Pyroglutamate carboxylic acid.  
CC FT CHAIN 1 1  
CC FT PEPTIDE 1 10  
CC FT PEPTIDE 14 >63  
CC FT ACT\_SITE 3 3  
CC FT MOD\_RES 1 1  
CC FT MOD\_RES 10 10  
CC FT NON\_TER 63  
CC SO SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;

Query Match 35.1%; Score 52; DB 1; Length 63;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 HMSGGLRP 27  
| | | | | | | |  
DB 2 HMSGGLRP 9

RESULT	9		
GONL_XENTLA	STANDARD;	PRT;	89 AA.
ID_GONI_XENTLA			
AC_P45656;			
DT_01-NOV-1995 (Rel. 32, Created)			
DT_01-NOV-1995 (Rel. 32, Last sequence update)			
DT_28-FEB-2003 (Rel. 41, Last annotation update)			
DE_Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH) (Lulliberin I).			
OS_Xenopus laevis (African clawed frog).			
OC_Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC_Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;			
OC_Xenopodinae; Xenopus.			
OX_NCBI_TaxID=8355;			
RN [1]			
RP_SEQUENCE FROM N.A.			
RC_TISSUE=Forebrain;			
RX_MEDLINE=94185563; PubMed=8137750;			
RA_Hayes W.P., Gray S., Batey J.F.;			
RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a			
RT mammalian-like expression pattern and conserved domains in			
RT GNRH-associated peptide, but brain onset is delayed until			
RT metamorphosis.";			
RL_Endocrinology 134:1835-1844(1994).			
CC -I FUNCTION: stimulates the secretion of gonadotrophins.			
CC -I SUBCELLULAR LOCATION: Secreted.			
CC -I SIMILARITY: Belongs to the GNRH family.			
-----			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way used			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
-----			
DR_EMBL: L28040; AAA49728.1; -.			
DR_PIR: I51423; I51423.			
DR_InterPro: IPRO02012; GNRH.			
DR_InterPro: IPRO04079; GonadoliberinI.			
DR Pfam: PF00446; GNRH; 1.			
DR PRINTS: PRO1541; GONADOLIBRNT.			
DR_PROSITE: PS00473; GNRH; 1.			
KW_Cleaveage on pair of basic residues; Hormone; Amidation; Hypothalamus;			
KM_Signal; Pyroglutamate carboxylic acid.			
FT SIGNAL	1	23	
FT CHAIN	24	89	PROGONADOLIBERIN I.
FT PEPTIDE	24	33	GONADOLIBERIN I.
FT PEPTIDE	37	89	GONADOTROPIN-RELEASING HORMONE ASSOCIATED PEPTIDE.
FT MOD_RES	37	85	GNRH-ASSOCIATED PEPTIDE I (GAP).
FT MOD_RES	33	24	PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES	33	24	AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ_SEQUENCE	89 AA:	10246 MM;	6FAFE36FBAED0D284 CRC64;
Query Match		35.1%;	Score 52; DB 1; Length 89;
Best Local Similarity		100.0%;	Pred. No. 0.56;
Matches	8; Conservative	0; Mismatches	0; Indels
			Gaps
OY	20 HMSYGLRP	27	
Dd			
	25 HMSYGLRP	32	
RESULT	10		
GONI_TUPGB	STANDARD;	PRT;	92 AA.
ID_GONI_TUPGB			
AC_Q95335;			
DT_15-DEC-1998 (Rel. 37, Created)			
DT_15-DEC-1998 (Rel. 37, Last sequence update)			
DT_28-FEB-2003 (Rel. 41, Last annotation update)			
DE_Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)]			

```

DE (uteinizing hormone-releasing hormone 1) (gonadotrophin-releasing
DE hormone 1) (GNRH 1) (luliberin 1); GNRH-associated peptide 1].
GN GNRH1 OR GNRH.
OS Tupaiia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxId:9396;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal.";
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HOMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the GNRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see license@isb-sib.ch).
CC -----
CC DR EMBL; U63326; AAB16837.1; -.
CC DR InterPro; IPR002012; GNRH.
CC DR InterPro; IPR004079; Gonadoliblerin1.
CC DR Pfam; PF00046; GNRH; 1.
CC DR PRINTS; PR01541; GONADOLIBRIN1.
CC DR PROSITE; PS00473; GNRH; 1.
CC KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Placental; Signal; Pyroglutamate carboxylic acid.
CC FT FT SIGNAL 1 23 BY SIMILARITY.
CC FT CHAIN 24 92 GONADOLIBERIN 1.
CC FT PEPTIDE 24 33 GONADOLIBERIN I.
CC FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE 1.
CC FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC ACTIVITY.
CC FT MOD_RES 24 24 PYROGLUTAMATE CARBOXYLIC ACID (BY
CC SIMILARITY).
CC FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
CC SIMILARITY).
CC SO SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;
CC -----
Query Match 35.1%; Score 52; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 20 HWSYGLRP 27
Db 25 HWSYGLRP 32
-----
RESULT 11
YDQ4_SCHPO
ID YDQ4_SCHPO STANDARD; PRT; 452 AA.
AC Q14197;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C5D6.04 in chromosome 1.
GN SPAC5D6.04.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.

```

OX NCBI\_TaxID=4896;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 RA James K., Jones I., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odel C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,  
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller Auer S.,  
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Mambuti R., Purrelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Kochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpkovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RA "The genome sequence of Schizosaccharomyces pombe.";  
 RT Nature 415:871-880(2002).  
 RL Nature 415:871-880(2002).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: TO YEAST YBR287W.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: 298056; CAB10852.1; -  
 DR PIR: T38962; T38962.  
 DR GeneDB: Spombe; SPAC5D6.04; -  
 DR InterPro: IPR004776; Auxin\_eff.  
 DR Pfam: PF03547; Auxin\_eff.1.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 18 38  
 FT TRANSMEM 81 101 POTENTIAL.  
 FT TRANSMEM 269 289 POTENTIAL.  
 FT TRANSMEM 317 337 POTENTIAL.  
 FT TRANSMEM 354 374 POTENTIAL.  
 FT TRANSMEM 390 410 POTENTIAL.  
 FT TRANSMEM 428 448 POTENTIAL.  
 SQ SEQUENCE 452 AA: 49575 MW: 71877BA5725C69A8 CRC64:  
 Query Match 33.1% Score 49; DB 1; Length 452;  
 Best Local Similarity 57.1% Pred. No. 8.3;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 13 TELGSLHMSYGLR 26  
 DB 166 SQLGALRMSYGYR 179  
 RESULT 12  
 HMCCT\_BOMMO STANDARD; PRT; 3133 AA.  
 AC P98092;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Hemocytin precursor (Humoral lectin).  
 OS Bombyx mori (Silk moth).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
 OC Bombycidae; Bombyx.  
 OX NCBI\_TaxID=7091;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Fuyou X Tokai; TISSUE=Hemocyte;  
 RX MEDLINE=95178544; PubMed=7873598;  
 RA Kotani E., Yamakawa M., Iwamoto S.-I., Tashiro M., Mori H.,  
 RA Matsubara F., Tanai K., Kadoya-Okuda K., Kato Y., Mori H.;  
 RT "Cloning and expression of the gene of hemocytin, an insect humoral  
 RT lectin which is homologous with the mammalian von Willebrand  
 RT factor.";  
 RL Biochim. Biophys. Acta 1260:245-258(1995).  
 RN [2]  
 RP SEQUENCE OF 2221-3133 FROM N.A.  
 RA Kotani E., Iwamoto S.I., Tashiro M., Mori H., Sumida M.,  
 RA Matsubara F., Yamakawa M.;  
 RL Submitted (JAN-1993) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: ADHESIVE PROTEIN AND RELATES TO HEMOSTASIS OR  
 CC ENCAPSULATION OF FOREIGN SUBSTANCES FOR SELF-DEFENSE.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN HEMOCYTES DURING LARVAL-PUPAL  
 CC METAMORPHOSIS.  
 CC -1- INDUCTION: HEMAGGLUTINATION ACTIVITY IS INCREASED BY BACTERIAL  
 CC OR VIRAL INFECTION AND INHIBITED BY D-MANNOSE, N-ACETYL-D-  
 CC GALACTOSAMINE AND D-MALTOSE.  
 CC -1- PMW: MAY BE CONVERTED INTO THE 260 kDa MATURE HEMOCYTIN BY  
 CC PROTEOLYSIS.  
 CC -1- SIMILARITY: TO MAMMALIAN VON WILLEBRAND FACTOR.  
 CC -1- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT  
 CC OF HUMAN MUCIN 2.  
 CC -1- SIMILARITY: Contains 2 F5/8 type C domains.  
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D29738; BAA06160.1; -  
 DR EMBL: D14035; BAA03124.1; -  
 DR PIR: S52093; S52093.  
 DR HSPD: P12259; IG2R.  
 DR InterPro: IPR006207; Cys\_knot\_C.  
 DR InterPro: IPR000421; FA58\_C.  
 DR InterPro: IPR002172; LDL\_receptor\_A.  
 DR InterPro: IPR002919; TIL\_Cysrich.  
 DR InterPro: IPR001007; VWF\_C.  
 DR InterPro: IPR001846; VWF\_D.  
 DR Pfam: PF00734; F5\_F8\_type\_C; 2.  
 DR Pfam: PF01826; TIL; 6.  
 DR Pfam: PF00094; vwd; 3.  
 DR SMART: SM00231; FA58C; 2.  
 DR SMART: SM00192; LDla; 1.  
 DR SMART: SM00216; vwd; 3.  
 DR PROSITE: PS01185; CTCK\_1; 1.  
 DR PROSITE: PS01225; CTCK\_2; 1.  
 DR PROSITE: PS01285; FA58C\_1; 2.  
 DR PROSITE: PS01286; FA58C\_2; 2.  
 DR PROSITE: PS50022; FA58C\_3; 2.  
 DR PROSITE: PS01208; VWF\_C\_1; FALSE\_NEG.  
 DR Lectin; Glycoprotein; Signal; Repeat; Cell adhesion.  
 KW SIGNAL 1 3133  
 FT CHAIN 29 131  
 FT DOMAIN 153 240  
 FT DOMAIN 153 240

```

FT DOMAIN 248 613 D1.
FT DOMAIN 940 1095 F5/8 TYPE C 1.
FT DOMAIN 1116 1254 F5/8 TYPE C 2.
FT DOMAIN 1283 1356 D1.
FT DOMAIN 1620 1951 D2.
FT DOMAIN 1952 2315 D3.
FT DOMAIN 2230 2321 C-TYPE LECTIN (ATYPICAL).
FT DOMAIN 2235 2361 B1.
FT DOMAIN 2435 2469 B2.
FT DOMAIN 2553 2622 WFC 1.
FT DOMAIN 2842 2907 WFC 2.
FT DOMAIN 2971 3076 CTRK.
FT DOMAIN 895 914 POLY-THR.
FT DOMAIN 1267 1270 POLY-GLU.
FT DOMAIN 1425 1428 POLY-THR.
FT DOMAIN 1447 1450 POLY-THR.
FT DOMAIN 1474 1479 POLY-SER.
FT DOMAIN 2148 2153 POLY-PRO.
FT DOMAIN 2156 2159 POLY-PRO.
FT DOMAIN 2341 2344 POLY-PRO.
FT DISULFID 940 1095 BY SIMILARITY.
FT DISULFID 1116 1254 BY SIMILARITY.
FT DISULFID 2981 3040 BY SIMILARITY.
FT DISULFID 2991 3054 BY SIMILARITY.
FT DISULFID 3004 3070 BY SIMILARITY.
FT DISULFID 3020 3072 BY SIMILARITY.
FT DISULFID ? 3075 BY SIMILARITY.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1170 1170 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1387 1387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1622 1622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1727 1727 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1847 1847 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1975 1975 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1985 1985 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2093 2093 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2113 2113 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2161 2161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2276 2276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2451 2451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2647 2647 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2654 2654 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2663 2663 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2794 2794 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2810 2810 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2865 2865 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2929 2929 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2964 2964 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3028 3028 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 1288 1288 R -> G.
FT VARIANT 1305 1305 T -> S.
SQ SEQUENCE 3133 AA; 343350 MW; E5210D5D14A7B2B2 CRC64;

Query Match 33.1%; Score 49; DB 1; Length 3133;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McCreary J.E., Park M.,
RT Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
RL from brains of the American alligator (Alligator mississippiensis).";
CC Regul. Pept. 33:105-116(1991).
CC -1- FUNCTION: Stimulates the secretion of gonadotropins.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the GnRH family.
CC PIR: A60066; RHA01.
CC InterPro: IPR002012; GnRH.
CC Pfam: PF00446; GnRH; 1.
CC PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus; Pyroglutamate carboxylic acid.
FT MOD_RES 1 10 AMIDATION.
FT MOD_RES 10 10 PYROGLUTAMATE CARBOXYLIC ACID.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D728645A3 CRC64;

Query Match 32.4%; Score 48; DB 1; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.24;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 20 HMSGRLP 27
DB 2 HMSGRLP 9

RESULT 14
ID GONL_CAVPO STANDARD; PRT; 92 AA.
AC 054713;
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadolibertin I precursor [contains: Gonadolibertin I (LH-RH I)
DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
GN GNRI OR GNRI OR LHRI.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley white; TISSUE=Hypothalamus;
RX MEDLINE=97462693; PubMed=932920;
RA Jimenez-Linan M., Rubin B.S., King J.C.;
RT "Examination of guinea pig luteinizing hormone-releasing hormone gene
RT brain.";
RT reveals a unique decapeptide and existence of two transcripts in the
RT brain.";
RL Endocrinology 138:4123-4130(1997).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the GnRH family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL, AF033346; AAB87688.1; -

```

```

DR InterPro: IPR002012; GNRH.
DR InterPro: IPR004079; GonadolibertinI.
DR Pfam: PF00446; GNRH. 1.
DR PRINTS: PR01541; GONADOLIBRINT.
DR PROSITE: PS00473; GNRH. 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT MOD_RES 10279 MW: ACET74613F456D663 CRC64;
SQ SEQUENCE 92 AA: 10279 MW: ACET74613F456D663 CRC64;

Query Match 32.4%; Score 48; DB 1; Length 92;
Best Local Similarity 57.1%; Pred. No. 2.3;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 ELGSLMWSYGLRP 27
   | | : ||| : ||
Db 19 ENGSGQYWSYGVPR 32

RESULT 15
GONL_CHICK STANDARD; PRT: 92 AA.
AC P37042; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonaadolibertin I precursor [Contains: Gonadolibertin I (LH-RH I)
DE (lutetizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GNRH I) (Lulberin I); GNRH-associated peptide I].
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=White leghorn;
RX MEDLINE=94059355; PubMed=7902095;
RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;
RT "Characterization of the chicken preprogonadotrophin-releasing
RT hormone-I gene.";
RL J. Mol. Endocrinol. 11:19-29(1993).
RN [2]
RP SEQUENCE OF 24-33.
RC TISSUE=Hypothalamus;
RX MEDLINE=82265778; PubMed=7050119;
RA King J.A., Millar R.P.;
RT "Structure of chicken hypothalamic lutetizing hormone-releasing
RT hormone. II. Isolation and characterization.";
RL J. Biol. Chem. 257:10729-10732(1982).
RN [3]
RP SEQUENCE OF 24-33.
RC TISSUE=Hypothalamus;
RX MEDLINE=82265777; PubMed=7050118;
RA King J.A., Millar R.P.;
RT "Structure of chicken hypothalamic lutetizing hormone-releasing
RT hormone. I. Structural determination on partially purified
RT material.";
RL J. Biol. Chem. 257:10722-10728(1982).

```

```

CC -! FUNCTION: Stimulates the secretion of gonadotropins.
CC -! SUBCELLULAR LOCATION: Secreted.
CC -! SIMILARITY: Belongs to the GNRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X69491; CAA49246.1; -.
CC PIR: I50644; I50644.
CC InterPro: IPR002012; GNRH.
CC InterPro: IPR004079; GonadolibertinI.
CC Pfam: PF00446; GNRH. 1.
CC PRINTS: PR01541; GONADOLIBRINT.
CC PROSITE: PS00473; GNRH. 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT MOD_RES 24 24
FT MOD_RES 33 33
SQ SEQUENCE 92 AA: 10206 MW: 61AB7EBAF508B6A CRC64;

Query Match 32.4%; Score 48; DB 1; Length 92;
Best Local Similarity 87.5%; Pred. No. 2.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
   |||||:|
Db 25 HWSYGLRP 32

Search completed: September 24, 2003, 17:41:18
Job time : 6.29668 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:29:11 ; Search time 27.1468 seconds

(without alignments)  
266.163 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYKANSKFIGTELPSLHMSYGLRPX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database :
- 1: SPTREMBL\_23:\*
  - 2: sp\_archaea:\*
  - 3: sp\_bacteria:\*
  - 4: sp\_fungi:\*
  - 5: sp\_human:\*
  - 6: sp\_invertebrate:\*
  - 7: sp\_mammal:\*
  - 8: sp\_mhc:\*
  - 9: sp\_organelle:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_virus:\*
  - 16: sp\_bacteriap:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	50.0	1310	2 Q93N27	Q93N27 clostridium
2	52.5	35.5	342	16 Q98F20	Q98F20 rhizobium i
3	52	35.1	90	13 Q90Y63	Q90Y63 rana catesb
4	52	35.1	91	13 Q9PRH0	Q9PRH0 anguilla ja
5	52	35.1	226	12 Q8BAZ8	Q8BAZ8 hepatitis b
6	52	35.1	382	5 Q95VY3	Q95VY3 lumbricus t
7	51.5	34.8	834	16 Q8U7P1	Q8U7P1 agrobacteri
8	51	34.5	384	5 Q77072	Q77072 eisenia foe
9	51	34.5	388	17 Q9YD14	Q9YD14 aeropyrum p
10	51	34.5	524	4 Q8N8Z6	Q8N8Z6 homo sapien
11	51	34.5	539	4 Q81YK5	Q81YK5 homo sapien
12	50.5	34.1	60	7 Q31585	Q31585 salmo salar
13	50.5	34.1	71	7 Q9XRJ9	Q9XRJ9 salvelinus
14	50.5	34.1	85	7 Q95IS2	Q95IS2 salmo salar
15	50.5	34.1	85	7 Q95HY1	Q95HY1 salmo salar
16	50.5	34.1	86	7 Q95HX4	Q95HX4 salmo salar

## ALIGNMENTS

17	50.5	34.1	244	7 Q31590	Q31590 salmo salar
18	50	33.8	94	13 Q8UUK6	Q8UUK6 sclerophages
19	50	33.8	274	3 Q14264	Q14264 schizosacch
20	50	33.8	480	16 Q91482	Q91482 pseudomonas
21	49.5	33.4	67	7 Q31578	Q31578 salmo salar
22	49	33.1	103	12 Q8QSV9	Q8QSV9 hepatitis b
23	49	33.1	324	12 Q89914	Q89914 sonchus yel
24	49	33.1	484	16 Q9H230	Q9H230 pseudomonas
25	48	32.4	318	16 Q8UES9	Q8UES9 agrobacteri
26	48	32.4	794	16 Q8ZDY5	Q8ZDY5 yersinia pe
27	48	32.4	1036	16 Q8P1X2	Q8P1X2 xanthomonas
28	48	32.4	2091	3 P78616	P78616 emeritella
29	47.5	32.1	84	13 Q9DEK4	Q9DEK4 coregonus s
30	47.5	32.1	85	7 Q95IS3	Q95IS3 salmo salar
31	47.5	32.1	85	7 Q95IR2	Q95IR2 salmo salar
32	47.5	32.1	149	7 Q31495	Q31495 oncorhynch
33	47.5	32.1	205	17 Q981D4	Q981D4 sulfobolus
34	47.5	32.1	216	7 Q9G3H0	Q9G3H0 salmo trutt
35	47.5	32.1	216	7 Q9G3G9	Q9G3G9 salmo trutt
36	47.5	32.1	1361	12 Q8BF56	Q8BF56 tulip virus
37	47	31.8	114	10 Q818X1	Q818X1 arabidopsis
38	47	31.8	134	10 Q9LSN9	Q9LSN9 arabidopsis
39	47	31.8	230	8 Q8HMC7	Q8HMC7 physiculus
40	47	31.8	599	2 Q810X1	Q810X1 actinobact
41	47	31.8	728	16 Q9HYO6	Q9HYO6 pseudomonas
42	46.5	31.4	67	7 Q31582	Q31582 salmo salar
43	46.5	31.4	67	7 Q31577	Q31577 salmo salar
44	46.5	31.4	67	7 Q31581	Q31581 salmo salar
45	46.5	31.4	85	7 Q95HX8	Q95HX8 salmo salar

## RESULT 1

Q93N27	PRELIMINARY;	PRT;	1310 AA.
AC	Q93N27;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	Tetanus toxin (Fragment).		
OS	Clostridium tetani.		
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;		
OC	Clostridium.		
OX	NCBI_TaxID=1513;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Shumin Z., Dianliang L.;		
RT	"Cloning and sequence analysis of tetanus toxin gene.";		
RU	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF389424; AKK72964.2; "		
DR	InterPro; IPR000395; Bontoxilysin.		
DR	InterPro; IPR001064; Crystallin.		
DR	InterPro; IPR006025; zn_MTPeptide.		
DR	Pfam; PF01742; Peptidase_M27; 1.		
DR	PRINTS; PR00760; BONTOTOXILYSIN.		
DR	ProDom; PD001963; Bontoxilysin; 1.		
DR	PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.		
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.		
FT	NON_TER		
FT	NON_TER		
SO	SEQUENCE		
	1310 AA; 150316 MW; 9EADDC914418E450 CRC64;		
QY	Query Match	50.0%; Score 74; DB 2; Length 1310;	
	Best local Similarity	100.0%; Pred. No. 0.021;	
	Matches	15; Conservative	0; Mismatches
		0; Indels	0; Gaps
DB	1 QYKANSKFIGTEL 15		
	831 QYKANSKFIGTEL 845		

RESULT 2			
ID	Q98FZ0	PRELIMINARY;	PRT; 342 AA.
AC	Q98FZ0;		
DT	01-OCT-2001 (TREMBLrel. 18, Created)		
DT	01-OCT-2001 (TREMBLrel. 18, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	Hypothetical protein ml13560.		
GN	ML13560.		
OS	Rhizobium loti (Mesorhizobium loti).		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;		
OC	Phyllobacteriaceae; Mesorhizobium.		
OX	NCBI_TaxID=381;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MAEF30309;		
RX	MEDLINE=21082930; PubMed=11214968;		
RA	Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,		
RA	Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,		
RA	Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsuo A.,		
RA	Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,		
RA	Takeuchi C., Yamada M., Tabata S.;		
RT	"Complete genome structure of the nitrogen-fixing symbiotic bacterium		
RT	Mesorhizobium loti.";		
RL	DNA Res. 7:331-338(2000).		
DR	EMBL: AP003002; BAB50426.1; -;		
KW	Hypothetical protein; Complete proteome.		
QO	SEQUENCE 342 AA; 37582 MW; 99597D57D25D1A1 CRC64;		

Query Match	35.5%	Score 52.5	DB 16	Length 342
Best Local Similarity	37.5%	Pred. No. 9.7		
Matches	9	Conservative 7	Mismatches 5	Indels 3
			Gaps 1	
QY	1 QYIKANSKFGITELGDSLHWSYG 24			
	::: :   :::			
db	82 RFLKAGSDFLGADTG---YWFPG 102			

	RESULT 3
ID	090Y63
DT	090Y63
AC	PRELIMINARY;
PRT:	90 AA.
DT	01-DEC-2001 (TREMBLrel. 19, Created)
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE	Progonadoliberin I precursor [contains: Gonadolibrin I (LNHR I) (gonadotropin releasing hormone I), GnRH-associated peptide I (GAP1)].
GN	(GNRH1 OR GNRRH.
OS	Rana catesbeiana (Bull frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX	NCBI_Taxid=8400;
RN	[1]
RP	SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC	TISSUE=Forebrain;
RX	MEDLINE=21102951; PubMed=11170016;
RA	Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA	Kwon H.B.;
RT	"Cloning and characterization of cDNAs encoding the GNRH1 and GNRHZ precursors from bullfrog ( <i>Rana catesbeiana</i> ).";
RL	J. Exp. Zool. 289:190-201(2001).
CC	-!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY SIMILARITY).
CC	-!- SUBCELLULAR LOCATION: SECRETED.
CC	-!- TISSUE SPECIFICITY: FOREBRAIN.
CC	-!- DEVELOPMENTAL STAGE: EXPRESSED AT SIGNIFICANTLY HIGHER LEVELS DURING POST-BREEDING. NOT EXPRESSED IN PITUITARY.
CC	-!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
EMBL:	AF188754; AAL05972.1; ..
DR	Interpro: IPRO02012; GnRH.
DR	Interpro: IPRO04079; Gonadolibrini.

DR	PFam:	PF00446;	GNRH; 1.	
DR	PRINTS;	PRO154;	GONADOLIBRI.	
DR	PROSITE;	PS00473;	GNRH; 1.	
KW	Cleavage on pair of basic residues; Hormone; Amidation; Signal;			
KW	Pyrrolidone carboxylic acid.			
FT	SIGNAL	1	24	POTENTIAL.
FT	CHAIN	25	90	PROGONADOLIBERIN I.
FT	PEPTIDE	25	34	GONADOLIBERIN I.
FT	PEPTIDE	38	86	GNRH-ASSOCIATED PEPTIDE I
FT				(BY SIMILARITY).
FT	MOD_RES	25	25	PYRROLIDONE CARBOXYLIC ACID
FT				(BY SIMILARITY).
FT	MOD_RES	34	34	AMIDATION (G-35 PROVIDE AMIDE GROUP)
FT				(BY SIMILARITY).
FO	SEQUENCE	90 AA;	10291 MW;	317203B4E3DAFE7 CRC64;

Query Match	35.1%	Score 52	DB 13	Length 90
Best Local Similarity	100.0%	Pred. No. 2.5		
Matches	8	Conservative	0	Mismatches 0
				Indels 0
				Gaps 0
QY	20	HWSTGLRP	27	
db	26	HWSTGLRP	33	

RESULT 4  
Q9PRH0 PRELIMINARY; PRT; 91 AA  
ID Q9PRH0

```

DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE prepro-mGnRH precursor (Gonadoliberin) (Gonadotropin-releasing
OS hormone) (LH-RH) (Luliberin) .
OS Anguilla japonica (Japanese eel) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguilla.
OX NCBI_TaxID=7937;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RC Okubo K., Suetake H., Aida K.;
RT "Expression of two gonadotropin-releasing hormone (GnRH) precursor
RT genes in various tissues of the Japanese eel and evolution of GnRH." ;
RL Zool. Sci. 16:471-478(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC Okubo K., Suetake H., Aida K.;
RT "A splicing variant for the prepro-mammalian gonadotropin-releasing
RT hormone (prepro-mGnRH) mRNA is present in the brain and various
RT peripheral tissues of the Japanese eel." ;
RL Zool. Sci. 16:645-651(1999).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY) .
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC EMBL; AB026989; BAA82608.1; -
CC EMBL; AB026991; BAA83597.1; -
CC InterPro; IPR002012; GnRH.
CC InterPro; IPR004079; Gonadoliberin1.
CC Pfam; PF00446; GnRH; 1
DR PRINTS; PR01541; GONADOLIBRIN1.
DR PROSITE; PS00473; GnRH; 1.
DR Amidaation; Hormone; Signal.
FT SIGNAL 1 22
FT CHAIN 23 32 POTENTIAL.
FT CHAIN 33 91 MGNRH.
FT CHAIN 91 91 GnRH ASSOCIATED PEPTIDE.
SQ SEQUENCE 91 AA, 9893 MW, BA15C9DC08434A7B CRC64;

Query Match 35.1%; Score 52; DB 13; Length 91;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 8; Conservative 0; Indels 0; Gaps 0;

```

OY 20 HWSYGLRP 27  
 |||||  
 DB 24 HWSYGLRP 31

## RESULT 5

08BA28 PRELIMINARY; PRT: 226 AA.  
 AC 08BA28:  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Surface antigen.  
 OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxId=10407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=F078;  
 RA Bowden R.J.;  
 RT "Hepatitis B virus variability and human population history in the Pacific";  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=F078;  
 RA Basuni A.A., Bowden R.J., Butterworth L.-A., Cooksley G., Locarnini S., Carman W.F.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY122488; AAM82819.1;  
 SQ SEQUENCE 226 AA; 63791C8CAEFFB514 CRC64;

Query Match 35.1%; Score 52; DB 12; Length 226;  
 Best Local Similarity 39.1%; Pred. No. 7.2;  
 Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 5 ANSKFIFTELGPSSLMWSYGLRP 27  
 |::|::|::|::|::|  
 DB 166 ASVRSWLSLAPFVHMFVGLSP 188

## RESULT 6

095VY3 PRELIMINARY; PRT: 382 AA.  
 AC 095VY3:  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Coelomic cytolitic factor precursor.  
 OS Lumbricus terrestris (Common earthworm).  
 OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;  
 OX Lumbricina; Lumbricidae; Lumbricus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Beschin A., De Baetselier P., Billef M.;  
 RT "Distinct carbohydrate recognition domains of an earthworm defense molecule recognize Gram negative and Gram positive bacteria";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF395805; AAL09587.1;  
 DR InterPro: IPR000757; Glyco\_hydro\_16.  
 DR Pfam: PF00722; Glyco\_hydro\_16; 1.  
 KW Signal.  
 FT SIGNAL.  
 SQ SEQUENCE 382 AA; 43931 MW; 5256CF171EB7D3FB CRC64;

Query Match 35.1%; Score 52; DB 5; Length 382;  
 Best Local Similarity 36.8%; Pred. No. 13;  
 Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

OY 6 NSKFIGITELGPSSLMWSYGLRP 24  
 ::::|::|::|::|::|  
 DB 196 SGEFLGIQKMGSTHMGPG 214

## RESULT 7

08U7P1 PRELIMINARY; PRT: 834 AA.  
 AC 08U7P1:  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Periplasmic nitrate reductase large subunit.  
 GN NAPA OR ATG4408 OR AGR\_L\_917.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OX Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxId=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-21608550; PubMed-11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Woo L., Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L., Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyee D. Sr., Chapman P., Clendenning J., Deatherage G., Gillet M., Grant C., Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A., Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D., Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M., Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M., Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V., Nester E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58.";  
 RL Science 294:2317-2323(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-21608551; PubMed-11743194;  
 RA Goodner B., Hinkley G., Gattung S., Miller N., Blanchard M., Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L., Hounell K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F., RA Mollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B., RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G., RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58.";  
 RL Science 294:2323-2328(2001).  
 DR EMBL; AE009369; AAL45202.1; ALT\_INIT.  
 DR EMBL; AE008245; AAK89031.1;  
 DR InterPro: IPR006656; Molybdopterin.  
 DR InterPro: IPR006963; Molybdopterin\_FeS4.  
 DR InterPro: IPR006657; Mol\_dinuc\_bind.  
 DR InterPro: IPR006655; Prok\_Mboxred.  
 DR InterPro: IPR006311; Tat.  
 DR Pfam: PF00384; molybdopterin; 1.  
 DR Pfam: PF04879; Molybdop\_FeS4; 1.  
 DR Pfam: PF01568; Molybdop\_binding; 1.  
 DR TIGRFAMS; TIGR01409; Tat\_signal\_seq; 1.  
 DR PROSITE; PS00551; MOLYBDOPTERIN\_PROK\_1; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 834 AA; 93527 MW; FF707BC71CA08B9E CRC64;

Query Match 34.8%; Score 51.5; DB 16; Length 834;  
 Best Local Similarity 37.0%; Pred. No. 38;  
 Matches 10; Conservative 9; Mismatches 1; Indels 7; Gaps 2;

OY 2 YIKANSKFI-GITELGPSSLMWSYGLRP 27  
 ::::|::|::|::|::|  
 DB 289 FVRNHTKFRGVTDIG-----YGLRP 309

## RESULT 8

077072 PRELIMINARY; PRT: 384 AA.  
 AC 077072:  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

Query Match	34.5%;	Score 51;	DB 17;	Length 388;
Best Local Similarity	71.4%;	Pred. No. 19;		
Matches	10;	Conservative	3;	Indels 0; Gaps 0;
QY	14	ELGPSLHMSYGLRP	27	

Query Match	34.58;	Score 51;	DB 4;	Length 539;
Best Local Similarity	40.98;	Pred. No. 28;		

```

Best Local Similarity  57.9%;  Pred. No. 3.3;
Matches  11;  Conservative  3;  Mismatches  2;  Indels  3;  Gaps  1.;

QY      1 QYIKANS--KFGIGTEIG 16
      :::::  ||: ||||

```

095152

0951S2		
ID	0951S2	PRELIMINARY; PRT; 85 AA.
AC	0951S2;	
DT	01-DEC-2001 (TREMBLrel, 19, Created)	
DT	01-DEC-2001 (TREMBLrel, 19, last sequence update)	
DT	01-MAR-2002 (TREMBLrel, 20, last annotation update)	
DE	MHC class II beta chain (Fragment).	
OS	Salmo salar (Atlantic salmon).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;	
OC	Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.	
XC	NCBI TaxId=8030;	

RI  
PT

RT geographic scales at Major Histocompatibility Complex and  
 RT microsatellite in Atlantic salmon (Salmo salar)";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF336593; AK61882.1;  
 DR InterPro: IPR000353; MHC\_II\_beta.  
 DR Pfam: PF00965; MHC\_II\_beta; 1.  
 DR ProDom: PD000328; MHC\_II\_beta; 1.  
 KW Glycoprotein; MHC; MHC II; Transmembrane.  
 KW NON\_TER 1

## Query M

SO	SEQUENCE	85	AA:	9743	MW:	3214E01AD1BB66AC5	CRC64:	
	Query Match		34.1%:	Score 50.5:	DB 7:	Length 85:		
	Best Local Similarity		57.9%:	Pred. No. 4.1:				
	Matches 11:	Conservative		3:	Mismatches		2:	Indels 3:
								Gaps 1:
QY	1 QYIKANS---KFIGITEIG	16						

## RESULT 15

RESULT	15		
095HY1			
ID	095HY1	PRELIMINARY;	PRT; 85 AA.
AC	095HY1;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	MHC class II B antigen (Frigment).		
GN	DBI.		
OS	Salmo salar (Atlantic salmon).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;		

RX  
MEDI

RP SEQUENCE FROM N.A.  
RX MEDLINE=21383619; PubMed=11491536;  
RA Langefors A., Lohm J., von Schantz T.  
RT "Allelic polymorphism in MHC class II B in four populations of  
RT Atlantic salmon (Salmo salar)."  
RL Immunogenetics 53:329-336(2001).  
EMBL: AF104370; AL04002.1; -  
DR InterPro: IPR000353; MHC\_II\_beta.  
DR Pfam: PF00969; MHC\_II\_beta.1.  
DR ProDom: PD000328; MHC\_II\_beta.1.  
DR

34.18; Score 50.5; DB 7; Length 71;

Thu Sep 25 08:58:53 2003

us-09-848-834a-11.rspt

Page 6

FT	NON_TER	85	85
SQ	SEQUENCE	85 AA;	9723 MW; 27C9F7931F1F01C4 CRC64;

Query Match	34.1%	Score 50.5;	DB 7;	Length 85;
Best Local Similarity	57.9%	Pred. No. 4.1;		
Matches 11; Conservative	3;	Mismatches 2;	Indels 3;	Gaps 1;

```
QY      1 QYIKANS--KFIGITELG 16
          :||: || ||:| ||||
Db      33 EYIRFNSTVGKFEVGYTELG 51
```

```
Search completed: September 24, 2003, 17:47:18
Job time : 29.1468 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

Run on: September 24, 2003, 16:58:36 ; Search time 37.2964 Seconds  
(without alignments)  
140.442 Million cell updates/sec

Scoring table: BLOSUM62 , Gapop 10.0 , Gapext 0.5

```
Searched:      1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters:  1107863
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database

1:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1980.DAT *
2:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1981.DAT *
3:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1982.DAT *
4:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1983.DAT *
5:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1984.DAT *
6:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1985.DAT *
7:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1986.DAT *
8:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1987.DAT *
9:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1988.DAT *
10:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1989.DAT *
11:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1990.DAT *
12:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1991.DAT *
13:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1992.DAT *
14:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1993.DAT *
15:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1994.DAT *
16:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1995.DAT *
17:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1996.DAT *
18:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1997.DAT *
19:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1998.DAT *
20:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA1999.DAT *
21:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA2000.DAT *
22:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA2001.DAT *
23:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA2002.DAT *
24:	/SIDSL/gcgdata/genseq/genseqp-emb1/AA2003.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	166	99.4	33	23	AAU11423	Synthetic immunoge	
2	157	92.9	51	23	AAU11431	Synthetic immunoge	
3	134.5	79.6	33	15	AAK62715	LHRH-containing im	
4	95	56.2	20	23	AAU11414	P. falciparum circ	
5	95	56.2	33	16	AAK83570	IGE CH4 region com	
6	95	56.2	36	23	AAU11427	Synthetic immunoge	
7	89	52.7	19	22	AAK98951	Vaccine related MH	
8	89	52.7	21	10	AAK91504	Sequence of modifi	
9	89	52.7	21	16	AAK82586	Plasmodium falcipa	

10	89	52.7	21	16	AA879920	Malaria circumsporozoite
11	89	52.7	21	16	AA875955	P. falciparum CS p
12	89	52.7	21	16	AA870912	Malaria circumsporozoite
13	89	52.7	21	17	AAW05612	Circumsporozoite h
14	89	52.7	21	18	AAW35440	T-cell stimulatory
15	89	52.7	21	20	AAAY23252	peptide derived fr
16	89	52.7	21	21	AAAY80071	Pathogen derived T
17	89	52.7	21	21	AAAY54553	T helper cell (Th)
18	89	52.7	21	21	AAAY58777	unidentified pepit
19	89	52.7	21	22	AA899706	Plasmodium falcipar
20	89	52.7	21	22	AA624428	Plasmodium falcipar
21	89	52.7	21	22	AA684517	Plasmodium falcipar
22	89	52.7	21	22	AA6882659	Plasmodium falcipar
23	89	52.7	21	22	AA6893666	Plasmodium falcipar
24	89	52.7	21	22	AA884447	Plasmodium falcipar
25	89	52.7	21	22	AA898457	Sequence of r help
26	89	52.7	21	23	ABJ16181	Plasmodium falcipar
27	89	52.7	21	23	ABJ11372	Zinc transporter p
28	89	52.7	21	23	ABJ05781	P falciparum HLA b
29	89	52.7	21	23	AB878050	p falciparum circur
30	89	52.7	21	23	ABJ01952	Loosely MHC -restri
31	89	52.7	21	23	ABP51501	158b1d7 related HL
32	89	52.7	21	23	ABG34858	Malaria circumspor
33	89	52.7	21	23	AAU91545	Plasmodium circr
34	89	52.7	21	23	AAU953659	P. falciparum circ
35	89	52.7	21	23	AA8944649	Plasmodium circums
36	89	52.7	21	24	ABR01863	Plasmodium falcipar
37	89	52.7	22	24	AA882077	Circumsporozoite (
38	89	52.7	33	22	AA683663	Malaria C5T3 prote
39	89	52.7	33	22	AA663516	Peptide comprising
40	89	52.7	143	21	AAV49252	a peptide which ma
41	89	52.7	218	21	AAV49253	N6 polypeptide car
42	89	52.7	240	21	AAV49254	N10 polypeptide ca
43	89	52.7	390	21	AAV49255	N11 polypeptide ca
44	89	50.9	424	14	AA837756	N19 polypeptide ca
45	86	50.9	424	14	AA837757	RTS protein. Syn

## ALIGNMENTS

RESULT 1
AAU011423
AAU011423 standard; peptide; 33 AA.
AAU011423:
12-MAR-2002 (first entry)
Synthetic immunogen peptide 4.
Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
luteinising hormone releasing hormone; LHRR; contraceptive;
promiscuous helper T-cell peptide epitope; immunomimic peptide epitope
breast cancer; uterine cancer; gynaecological cancer; endometriosis;
uterine fibroid; benign prostatic hypertrophy; prostate cancer.
Chimeric - Plasmodium falciparum.
Chimeric - Mammalia.
Synthetic.
Key
Location/Qualifiers
Peptide 1..20
/note= "Malaria CSP protein (378-398 aa)"
Peptide 21..24
/note= "Spacer peptide"
Peptide 25..33
/note= "Gonadotrophin releasing hormone epitope"
Modified-site 33
/note= "Amidated glycine or glycinamide"
WO200185763-A2.

PD 15-NOV-2001.  
 XX 04-MAY-2001; 2001WO-US14363.  
 XX 05-MAY-2000; 2000US-202328P.  
 XX (APHT-) APHTON CORP.  
 PA Grimes S, Michaeli D, Stevens VC;  
 PI WPI: 2002-049440/06.  
 DR  
 XX  
 PT Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue  
 XX  
 PS Claim 11; Page 8; 43pp; English.  
 XX  
 CC The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 CC  
 XX  
 SQ Sequence 33 AA:  
 Query Match 99.4%; Score 168; DB 23; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 2,3e-18;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DEKRIAMKAKSSVFVNVSGLHWSYGLRP 32  
 Db 1 DEKRIAMKAKSSVFVNVSGLHWSYGLRP 32  
 RESULT 2  
 ID AA011431 standard; peptide; 51 AA.  
 XX  
 AC AA011431;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Synthetic immunogen peptide 12.  
 XX  
 KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Plasmodium falciparum.  
 OS Chimeric - Mammalia.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT MISC-difference 1  
 FT /label= OTHER  
 FT /note= "Other- Pyro-glutamic acid or 5-oxo proline"  
 FT Peptide 1..10  
 FT /note= "Gonadotropin releasing hormone epitope  
 FT (1..10 aa)"  
 FT Peptide 11..16  
 FT /note= "Spacer peptide"  
 FT Peptide 17..36  
 FT  
 FH

FT /note= "Malaria CSP protein (378-398 aa)"  
 FT Peptide 37..42  
 FT /note= "Spacer peptide"  
 FT Peptide 43..51  
 FT /note= "Gonadotropin releasing hormone epitope  
 FT (2-10 aa)"  
 FT Modified-site 51  
 FT /note= "Amidated glycine or glycylamide"  
 XX  
 XX WO200185763-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 XX 04-MAY-2001; 2001WO-US14363.  
 XX 05-MAY-2000; 2000US-202328P.  
 XX  
 XX (APHT-) APHTON CORP.  
 PA Grimes S, Michaeli D, Stevens VC;  
 PI WPI: 2002-049440/06.  
 DR  
 XX  
 CC The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 CC  
 XX  
 SQ Sequence 51 AA:  
 Query Match 92.9%; Score 157; DB 23; Length 51;  
 Best Local Similarity 94.1%; Pred. No. 1.9e-16;  
 Matches 32; Conservative 0; Mismatches 0; Indels 2; Gaps 1;  
 QY 1 DEKRIAMKAKSSVFVNVSGLHWSYGLRP 32  
 Db 17 DEKRIAMKAKSSVFVNVSGLHWSYGLRP 50  
 RESULT 3  
 ID AAR62715 standard; peptide; 33 AA.  
 XX  
 AC AAR62715;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 10-SEP-1995 (first entry)  
 XX  
 DE LHRH-containing immunogenic peptide.  
 XX  
 KW Helper T cell epitope; universal immune stimulator; invasion; hapten;  
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;  
 KW androgen-dependent carcinoma; antitumour; infertility;  
 KW Plasmodium falciparum circumsporozoite.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers

XX	breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX	uterine fibroid; benign prostatic hypertrophy; prostate cancer;
XX	circumsporozoite protein; CSP.
XX	Plasmodium falciparum.
XX	MO200185763-A2.
XX	15-NOV-2001:
XX	04-MAY-2001; 2001WO-US14363.
XX	05-MAY-2000; 2000US-202328P.
XX	(APHT-) APHTON CORP.
XX	Grimes S, Michael D, Stevens VC;
XX	WPI; 2002-049440/06.
XX	Novel synthetic immunogen for inducing immune response against
XX	gonadotropin releasing hormone, comprises fusion peptide having
XX	promiscuous helper T-cell peptide epitope and immunomimic peptide
XX	epitope or its analogue
XX	-
XX	Disclosure: Page 28; 43pp; English.
XX	The invention relates to a synthetic immunogen for inducing specific
XX	antibodies against gonadotropin releasing hormone (GnRH also known
XX	as luteinising hormone releasing hormone, LHRH) comprising a fusion
XX	peptide which comprises a promiscuous helper T-cell peptide epitope and
XX	immunomimic peptide epitope or its analogue. The synthetic
XX	immunogen is useful inducing an immune response against GnRH in an
XX	animal subject, and as such is useful as a contraceptive and in the
XX	treatment of diseases such as cancer (of the breast, uterus and other
XX	gynaecological cancer), endometriosis, uterine fibroids, benign prostatic
XX	hypertrophy and prostate cancer. The immunogen is effective in eliciting
XX	high and specific anti-GnRH antibody titres. The present sequence
XX	is a peptide from circumsporozoite protein, CSP, a promiscuous helper
XX	T-cell peptide epitope used in the immunogen of the invention.
XX	Sequence 20 AA:
XX	Query Match 56.2%; Score 95; DB 23; Length 20;
XX	Best Local Similarity 100.0%; Pred. No. 1.8e-07;
XX	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	QY 1 DEKKIAKMKASVFNVYNS 20
XX	
XX	1 DEKKIAKMKASVFNVYNS 20
XX	Db
XX	RESULT 5
XX	AAR83570
XX	AAR83570 standard; peptide; 33 AA.
XX	AAR83570;
XX	13-JUN-1996 (first entry)
XX	IGF CH4 region contg. peptide immunogen for treating allergies.
XX	IGF: CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
XX	vaccine; allergy; antibody; constant heavy chain.
XX	Synthetic.
XX	MO9526365-A1.
XX	05-OCT-1995.
XX	24-MAR-1995; 95WO-US03741.

PR 25-OCT-1994; 94US-0328912.  
 PR 28-MAR-1994; 94US-0218461.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 PI Mang CY;  
 XX  
 DR WPI; 1995-351297/45.  
 XX  
 PT Synthetic peptide-based immunogen contg. Ige CH4 peptide and helper  
 PT T cell epitope - useful for eliciting antibody prodn. for allergy  
 PT treatment  
 PS  
 PS Claim 5; Page 72; 87pp; English.  
 XX  
 CC AAR82592-R82600 and AAR83560-R83581 are peptide immunogens that are  
 CC useful in vaccines for treating allergic reactions. In the immunogens,  
 CC an Ige CH4 peptide is attached C-terminally to a series of amino acids  
 CC including a helper T cell epitope. The immunogen may also opt. contain  
 CC a fatty acid or fatty acid derivative, an invasins domain or alpha-NH2.  
 CC The immunogen produces high titres of antibodies to the effector site  
 CC in human Ige heavy chain (the CH4 domain peptide) which inhibit mast  
 CC cell activation and reduce allergen-induced Ige prodn. The immunogens  
 CC may be used in either a radially branching multimeric form or a  
 CC linearly arranged monomeric form.  
 CC  
 XX  
 SQ Sequence 33 AA;  
 Query Match 56.2%; Score 95; DB 16; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2 EKKIAKMEKASSVFNVNSG 21  
 |||||  
 Db 3 EKKIAKMEKASSVFNVNSG 22  
 |||||  
 RESULT 6  
 ID AAU11427 standard; peptide; 36 AA.  
 XX  
 AC AAU11427;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Synthetic immunogen peptide 8.  
 XX  
 KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; Immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Plasmodium falciparum.  
 OS  
 OS Chimeric - Mammalia.  
 OS Synthetic.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1  
 FT /label- OTHER  
 FT /note- "Other- Pyro-glutamic acid or 5-oxo proline"  
 FT 1..10  
 FT /note- "Gonadotrophin releasing hormone epitope"  
 FT 11..16  
 FT /note- "Spacer peptide"  
 FT 17..36  
 FT /note- "Malaria CSP protein (378-398 aa)"  
 XX  
 XX WO200185763-A2.  
 XX  
 XX 15-NOV-2001.  
 XX  
 PD 04-MAY-2001; 2001WO-US14363.  
 PE

XX  
 PR 05-MAY-2000; 2000US-202328P.  
 XX  
 XX (APHT-) APHTON CORP.  
 XX  
 PA Grimes S, Michael D, Stevens VC;  
 XX  
 PI  
 XX  
 DR WPI; 2002-049440/06.  
 XX  
 PT Novel synthetic immunogen for inducing immune response against  
 PT gonadotrophin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue  
 PS  
 PS Claim 11; Page 10; 43pp; English.  
 XX  
 CC The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotrophin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 CC  
 XX  
 SQ Sequence 36 AA;  
 Query Match 56.2%; Score 95; DB 23; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 DEKRIAKMEKASSVFNVNS 20  
 |||||  
 Db 17 DEKRIAKMEKASSVFNVNS 36  
 |||||  
 RESULT 7  
 ID AAM98951 standard; Peptide; 19 AA.  
 XX  
 AC AAM98951;  
 XX  
 DT 07-DEC-2001 (first entry)  
 XX  
 DE Vaccine related MHC ligand peptide SEQ ID NO:54.  
 XX  
 KW Glutamic acid; glutamine; vaccine; major histocompatibility complex;  
 KW MHC; immunomodulator; antiallergic; endocrine; neuroprotectant;  
 KW virucidal; bactericidal; antiparasitic; fungicidal; cytostatic;  
 KW medicine; pharmaceutical; immune disorder; immune deficiency;  
 KW autoimmune; hypersensitivity; allergy; graft rejection; infection;  
 KW hormonal disorder; central nervous system disease; cancer; melanoma;  
 KW anti-melanoma vaccine; human immunodeficiency virus.  
 XX  
 OS Plasmodium malariae.  
 OS  
 OS  
 XX  
 XX WO200170772-A2.  
 XX  
 XX 27-SEP-2001.  
 XX  
 PD 22-MAR-2001; 2001WO-FR00872.  
 XX  
 PR 23-MAR-2000; 2000FR-0003711.  
 XX  
 XX (FABR ) FABRE MEDICAMENT SA PIERRE.  
 XX  
 PA Klingner-Hamouir C, Corvaia N, Beck A, Goetsch L;  
 XX  
 PI  
 XX  
 DR WPI; 2001-611470/70.

XX Stabilized pharmaceutical containing N-terminal glutamic acid or  
 PT glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt  
 PT with strong acid -

PS Claim 9; Page 39; 149pp; French.

XX The present invention describes a pharmaceutical compound (I) that  
 CC contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue  
 CC in the form of an addition salt with a strong, physiologically  
 CC acceptable acid (II). Also described are: (a) a pharmaceutical  
 CC composition containing at least one (I); (b) a vaccine containing at  
 CC least one (I) where this is a major histocompatibility complex (MHC)  
 CC ligand (Ia); (c) a method for in vitro diagnosis of diseases associated  
 CC with the presence of (Ia); (d) a kit for method (c) that includes a (Ia);  
 CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,  
 CC antiallergic, neuroprotectant, virucidal, bactericidal, antiparasitic,  
 CC fungicidal and cytostatic activities. (I) are useful, in human or  
 CC veterinary medicine, in pharmaceutical compositions (for treating immune  
 CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,  
 CC allergy, graft rejection, infection, hormonal disorders and central  
 CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in  
 CC vaccines for treatment or prevention of: (i) viral, bacterial, parasitic  
 CC or fungal infections; or (ii) of cancers. A particular application is in  
 CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of  
 CC diseases associated with interactions between MHC and (I), e.g. melanoma  
 CC and human immunodeficiency virus infection. AAM98898 to AAM99592  
 CC represent peptides which can be used in pharmaceutical compounds from  
 CC the present invention.

SQ Sequence 19 AA:

Query Match 52.7%; Score 89; DB 22; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIKMEKASSVFNVNS 20  
 |||||  
 DB 1 EKKIKMEKASSVFNVNS 19

RESULT 8  
 AAP91504

ID AAP91504 standard; peptide: 21 AA.

AC AAP91504;

DT 13-MAR-1992 (first entry)

DE Sequence of modified Plasmodium circumsporozoite T-cell epitope CS.T3.

XX Circumsporozoite peptide; T-cell epitope; immunogenic composition;

KW vaccine.

OS Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 1..2 /note= "May be H-Asp-Ile, H-Ile, or H-"

FT Misc-difference 19..21 /note= "May be Val-Asn-Ser-OH, Val-Asn-OH,

FT Val-OH or -OH "

XX EP343460-A.

XX 29-NOV-1989.

XX 12-MAY-1989; 89EP-0108618.

XX 24-MAY-1988; 88GB-0012214.

XX (HOFF ) HOFFMANN-LA ROCHE AG.

PI Sinigaglia F;

XX WPI; 1989-349561/48.

XX Modified plasmodium CS peptide - used as a universally recognised

PT T-cell epitope in vaccines to elicit an immune response against

XX pathogenic agents

PS Claim 1; Page 16; 23pp; English.

XX Peptide CS.T3 (residues 1-21) corresponds to residues 378-398 of the  
 CC CS protein from P. falciparum but contains 2 Ala residues in place  
 CC of the native protein's Cys residues at positions 384 and 389. Also  
 CC claimed is AAP91504 (or modified forms, see FT) associated with an  
 CC antigenic structure representing a B-cell epitope, pref. a multiple  
 CC antigenic peptide, esp. multimers of the repeat sequences NAMP  
 CC present in P. falciparum CS protein.

SQ Sequence 21 AA:

Query Match 52.7%; Score 89; DB 10; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIKMEKASSVFNVNS 20  
 |||||  
 DB 3 EKKIKMEKASSVFNVNS 21

RESULT 9  
 AAR82586

ID AAR82586 standard; peptide: 21 AA.

AC AAR82586;

DT 13-JUN-1996 (first entry)

DE Plasmodium falciparum circumsporozoite helper T cell epitope. PF.

XX Ige; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;

KW vaccine; allergy; antibody; constant heavy chain.

OS Plasmodium falciparum.

XX WO9526365-A1.

XX 05-OCT-1995.

XX 24-MAR-1995; 95WO-US03741.

XX 25-OCT-1994; 94US-0328912.

XX 28-MAR-1994; 94US-0218461.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 1995-351297/45.

XX Synthetic peptide-based immunogen conty. Ige CH4 peptide and helper  
 PT T cell epitope - useful for eliciting antibody prodn. for allergy  
 PT treatment

PS Claim 3; Page 23; 87pp; English.

XX AAR82571-91 are helper T cell epitopes which can be used in the  
 CC preparation of a peptide immunogen that is useful in vaccines for  
 CC treating allergic reactions. In the immunogen an Ige CH4 peptide  
 CC is attached C-terminally to a series of amino acids including a  
 CC helper T cell epitope. The immunogen may also opt. contain a fatty  
 CC acid or fatty acid derivative, an invasiv domain or alpha-NH2. The  
 CC immunogen produces high titres of antibodies to the effector site  
 CC in human Ige heavy chain (the CH4 domain peptide) which inhibit mast

CC cell activation and reduce allergen-induced Ige prodn. The immunogens  
CC may be used in either a radially branching multimeric form or a  
CC linearly arranged monomeric form.

XX Sequence 21 AA;

Query Match 52.7%; Score 89; DB 16; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIAKMEKASSVFNVNS 20  
DB 3 EKKIAKMEKASSVFNVNS 21

RESULT 10  
AAR78920  
ID AAR78920 standard; peptide; 21 AA.

XX AAR78920;

DT 25-MAR-2003 (updated)  
DT 27-MAR-1996 (first entry)

DE Malaria circumsporozoite 378-398 helper T lymphocyte epitope.

XX 378-398; cytotoxic T; CTL; epitope; helper T; HTL; cell;

KW lymphocyte; antigens; treatment; disease prevention; tumours;

XX viruses; parasites; malaria circumsporozoite; hepatitis B.

OS Malaria circumsporozoite.

XX WO9522317-A1.

XX 24-AUG-1995.

PF 16-FEB-1995; 95WO-US02121.

XX 16-FEB-1994; 94US-0197484.

XX (CYTE-) CYTEL CORP.

PI Vitellio MA, Chesnut RM, Sette AD, Celis E, Grey H;

XX WPI; 1995-302545/39.

PT Compn. inducing cytotoxic T lymphocyte response to pref. viral,  
PT bacterial, parasitic or tumour antigens - useful in the treatment  
PT and prevention of diseases associated with the antigen e.g.

XX hepatitis B

XX Disclosure; Page 29; 109pp; English.

CC A compn. which induces a cytotoxic T lymphocyte (CTL) response to  
CC an antigen (Ag) in a mammal comprises, a CTL Ag response inducing  
CC peptide and a lipid conjugated helper T cell inducing peptide (i.e.  
CC AAR78918-R78922). The compn. induces a CTL response to bacterial,  
CC viral or tumour Ags, and is therefore useful in the treatment and  
CC prevention of diseases associated with the Ag, e.g. hepatitis B.  
CC (Updated on 25-MAR-2003 to correct PI field.)

XX SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 16; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIAKMEKASSVFNVNS 20  
DB 3 EKKIAKMEKASSVFNVNS 21

RESULT 11

AAR75955  
ID AAR75955 standard; Peptide; 21 AA.

XX AAR75955;

DT 06-MAR-1996 (first entry)

DE P. falciparum CS protein epitope residues 378-398.

XX MAGE-3; melanoma antigen; vaccine; immune response; immunogenic peptide;

KW cytotoxic T lymphocyte response; CTL; melanoma; breast cancer; antibody.

XX Plasmodium falciparum.

OS WO9519783-A1.

XX 27-JUL-1995.

PF 25-JAN-1995; 95WO-US01000.

XX 25-JAN-1994; 94US-0186266.

XX (CYTE-) CYTEL CORP.

PI Celis E, Grey HM, Kubo RT, Sette A;

XX WPI; 1995-269270/35.

PT Immunogenic peptide(s) that induce immune response to cancer cells  
PT - that express a MAGE-3 protein peptide epitope used in vaccines or  
PT adoptive immuno:therapy to induce cytotoxic T lymphocytes

XX Disclosure; Page 14; 44pp; English.

CC AAR75942 is derived from the sequence of the melanoma antigen (MAGE-3)  
CC protein and can be used to elicit a primary cytotoxic T lymphocyte  
CC response against cells expressing MAGE-3. Synthetic peptides AAR75945-53  
CC can be used therapeutically to elicit CTL responses to melanoma, breast,  
CC colon, prostate, or other cells which express proteins with this  
CC epitope. The peptides have specific HLA-A1 binding capacity. The peptides  
CC can be also used in vaccines, esp. combined with peptides such as  
CC AAR75955-56, which are T-helper epitopes.

XX SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 16; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIAKMEKASSVFNVNS 20  
DB 3 EKKIAKMEKASSVFNVNS 21

RESULT 12

AAR70912  
ID AAR70912 standard; Protein; 21 AA.

XX AAR70912;

DT 25-MAR-2003 (updated)

DT 09-OCT-1995 (first entry)

DE Malaria circumsporozoite 378-398 T helper peptide.

XX Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;

XX Malaria circumsporozoite 378-398; T helper peptide.

OS Malaria circumsporozoite.

XX WO9504542-A1.

XX 16-FEB-1995.

PD

XX 02-AUG-1994; 94WO-US08721.  
 PF 06-AUG-1993; 93US-0103623.  
 XX (CYTE-) CYTEL CORP.  
 PA  
 PI Files JD, Livingston BD, Sette AD, Sidney JC;  
 XX WPI; 1995-090681/12.  
 DR  
 XX Human melanoma antigen, MAGE-1, peptide(s) - useful for  
 PT stimulating immune response against melanoma  
 XX  
 PS Disclosure: Page 13; 59pp; English.  
 CC The T helper peptides described in AAR70910-R70914 are used in  
 CC conjunction with the C-terminal MAGE-1 peptides described in  
 CC AAR70915 to AAR70969. Compsns. containing the T helper and MAGE-1  
 CC peptides can be administered, as a vaccine to patients susceptible  
 CC to MAGE associated tumours, e.g. melanomas.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 21 AA;  
 Query Match 52.7%; Score 89; DB 16; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 EKRIAKMEKASSVFNVNS 20  
 Db 3 EKRIAKMEKASSVFNVNS 21  
 RESULT 13  
 AAW05612  
 ID AAW05612 standard; peptide; 21 AA.  
 XX  
 AC AAW05612;  
 XX  
 DT 10-DEC-1996 (first entry)  
 XX  
 DE Circumsporozoite helper T cell epitope.  
 XX  
 KW Immunoglobulin; IgE; membrane protein; human; epsilon chain; hepatitis B;  
 KW membrane anchoring domain; helper T cell; surface antigen; core antigen;  
 KW pertussis toxin; tetanus toxin; measles virus F protein; immunotherapy;  
 KW Chlamydia trachomatis major outer membrane protein; immunogen; vaccine;  
 KW diphtheria toxin; plasmidium falciparum; circumsporozoite; E. coli T8af;  
 KW schistosoma mansoni; triose phosphate isomerase; allergic reaction;  
 KW allergic rhinitis; food allergy; anaphylaxis; virally-induced asthma;  
 KW antihistamine; decongestant; beta-2 agonist; immunosuppression;  
 KW corticosteroid.  
 XX  
 XX Plasmidium falciparum.  
 OS  
 XX W09612740-A1.  
 PN  
 PD 02-MAY-1996.  
 XX  
 PF 25-OCT-1995; 95WO-US13841.  
 XX  
 PR 25-OCT-1994; 94US-0328519.  
 XX  
 XX (UNBI-) UNITED BIOMEDICAL INC.  
 PA  
 XX Walfeld AM, Wang CY;  
 PI  
 XX WPI; 1996-230555/23.  
 DR  
 XX Peptide immunogen useful in treatment of allergy - comprises  
 PT membrane-bound IgE epsilon-chain peptide synthesised linearly in  
 PT tandem with T helper epitope peptide

XX Claim 2; Page 19; 53pp; English.  
 PS  
 CC AAW05957-W05616 represent helper T cell epitopes used in the peptide  
 CC immunogens of the invention. This sequence represents the Plasmidium  
 CC falciparum circumsporozoite helper T cell antigen. The peptides of the  
 CC invention contain one of these sequences, and a membrane-bound  
 CC immunoglobulin E (IgE) fragment (see AAW05595 and AAW05596). The  
 CC peptide immunogens of the invention can be used in vaccines for the  
 CC immunotherapeutic treatment of allergic reactions, including allergic  
 CC rhinitis, food allergies, anaphylaxis, or virally-induced asthma. The  
 CC immunogens overcome the short effective period of antihistamines,  
 CC decongestants, and beta-2 agonists, while preventing the broad  
 CC immunosuppression of corticosteroids. The peptides do not have the  
 CC potential side effects of restlessness or sedation (associated with  
 CC antihistamines), associated increased morbidity in asthmatics (as seen  
 CC with beta-2 agonists) and adverse hormonal activities (observed in  
 CC corticosteroid users).  
 CC  
 XX  
 SQ Sequence 21 AA;  
 Query Match 52.7%; Score 89; DB 17; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 EKRIAKMEKASSVFNVNS 20  
 Db 3 EKRIAKMEKASSVFNVNS 21  
 RESULT 14  
 AAW35440  
 ID AAW35440 standard; peptide; 21 AA.  
 XX  
 AC AAW35440;  
 XX  
 DT 25-MAR-2003 (updated)  
 XX  
 DT 22-APR-1998 (first entry)  
 XX  
 DE T-cell stimulatory peptide from Plasmidium falciparum.  
 XX  
 KW T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;  
 KW scaffold; inhibition; metastasis; wound healing; solid phase.  
 XX  
 OS Plasmidium falciparum.  
 XX  
 PN W09738011-A1.  
 XX  
 PD 16-OCT-1997.  
 XX  
 PF 03-APR-1997; 97WO-DK00146.  
 XX  
 PR 03-APR-1996; 96DK-0000398.  
 XX  
 XX (PEPR-) PEPRSEARCH AS.  
 PA  
 XX Heegaard PMH, Jakobsen PH;  
 PI  
 XX WPI; 1997-512645/47.  
 DR  
 XX Non-dendritic peptide carrier linked to a solid phase - useful as a  
 PT diagnostic agent and as a scaffold for production of chemical  
 PT derivatives  
 XX  
 PS Claim 30; Page 199; 262pp; English.  
 XX  
 CC A non-dendritic peptide carrier (A) has been developed which is coupled  
 CC through a linker to a solid phase, forming a complex of (A)-solid phase.  
 CC Where (A) comprises 10-50 amino acids capable of forming a secondary  
 CC structure in a benign buffer after liberation from the solid phase, and  
 CC further the (A)-solid phase complex comprises an immunogenic substance  
 CC and/or an immune mediator coupled on (A). The present sequence  
 CC represents a specifically claimed T-cell stimulatory peptide from the



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:46:20 ; Search time 9.78116 Seconds  
(without alignments)  
142.750 Million cell updates/sec

Title: US-09-848-834a-12  
Perfect score: 169  
Sequence: 1 DEKIKAKMEKASSVFNVNSGSPSLHMSYGLRPX 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents-AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	134.5	79.6	33	1 US-08-446-692-27	Sequence 27, Appl
2	134.5	79.6	33	2 US-08-488-351A-27	Sequence 27, Appl
3	89	52.7	21	1 US-08-186-266-6	Sequence 6, Appl
4	89	52.7	21	1 US-08-446-692-48	Sequence 48, Appl
5	89	52.7	21	2 US-08-488-351A-48	Sequence 48, Appl
6	89	52.7	21	3 US-09-100-409A-54	Sequence 54, Appl
7	89	52.7	21	4 US-08-464-496-17	Sequence 12, Appl
8	89	52.7	21	4 US-08-788-822A-12	Sequence 12, Appl
9	89	52.7	21	5 US-08-197-484-97	Sequence 97, Appl
10	89	52.7	21	5 PCT-US95-02121-97	Sequence 97, Appl
11	89	52.7	21	5 PCT-US95-13841-20	Sequence 20, Appl
12	86	50.9	423	2 US-08-760-797A-1	Sequence 1, Appl
13	86	50.9	424	2 US-08-760-797A-3	Sequence 1, Appl
14	86	50.9	424	3 US-08-932-929B-3	Sequence 1, Appl
15	86	50.9	424	3 US-08-932-929B-1	Sequence 1, Appl
16	82	48.5	412	1 US-08-313-288B-18	Sequence 18, Appl
17	79	46.7	17	4 US-08-464-496-16	Sequence 16, Appl
18	79	46.7	17	4 US-08-197-484-96	Sequence 96, Appl
19	79	46.7	17	5 PCT-US95-02121-96	Sequence 96, Appl
20	75	44.4	16	2 US-08-817-933A-7	Sequence 7, Appl
21	73.5	43.5	20	1 US-08-465-167A-20	Sequence 20, Appl
22	73.5	43.5	20	4 US-08-627-820-20	Sequence 20, Appl
23	73.5	43.5	20	5 PCT-US92-07218-17	Sequence 17, Appl
24	70	41.4	15	6 5169933-30	Sequence 17, Appl
25	66	39.1	30	3 US-09-100-414B-71	Sequence 71, Appl
26	66	39.1	30	3 US-09-303-323-71	Sequence 71, Appl
27	66	39.1	30	4 US-09-770-014-71	Sequence 71, Appl

28	65	38.5	30	3 US-09-100-414B-73	Sequence 73, Appl
29	65	38.5	30	3 US-09-303-323-73	Sequence 73, Appl
30	65	38.5	30	4 US-09-770-014-73	Sequence 73, Appl
31	63.5	37.6	16	1 US-08-465-167A-19	Sequence 19, Appl
32	63.5	37.6	16	4 US-08-627-820-19	Sequence 19, Appl
33	63.5	37.6	16	5 PCT-US92-07218-16	Sequence 16, Appl
34	62	36.1	21	1 US-08-305-871A-12	Sequence 12, Appl
35	61	36.1	30	3 US-09-100-414B-68	Sequence 68, Appl
36	61	36.1	30	3 US-09-303-323-68	Sequence 68, Appl
37	61	36.1	30	4 US-09-770-014-68	Sequence 68, Appl
38	59	34.9	49	1 US-08-387-156-4	Sequence 4, Appl
39	59	34.9	49	2 US-08-694-865-4	Sequence 4, Appl
40	59	34.9	49	2 US-08-878-748-4	Sequence 4, Appl
41	59	34.9	49	3 US-09-124-491-4	Sequence 4, Appl
42	59	34.9	49	4 US-09-383-912-4	Sequence 4, Appl
43	59	34.9	544	1 US-08-387-156-10	Sequence 10, Appl
44	59	34.9	544	2 US-08-694-865-10	Sequence 10, Appl
45	59	34.9	544	2 US-08-878-748-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-08-446-692-27  
; Sequence 27, Application US/08446692  
; Patent No. 5759551  
; GENERAL INFORMATION:  
; APPLICANT: Ladd, Anna  
; APPLICANT: Wang, Chang YI  
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maria C.H. Lin  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: US  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,692  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria C.H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)415-8745  
; TELEFAX: (516)751-6849  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 33 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-446-692-27

Query Match 79.6%; Score 134.5; DB 1; Length 33;  
Best local Similarity 90.3%; Pred. No. 2.7e-13;  
Matches 28; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2 EKIKAKMEKASSVFNVNSGSPSLHMSYGLRP 32  
DB 3 EKIKAKMEKASSVFNVNSGSPSLHMSYGLRP 32

RESULT 2  
US-08-488-351A-27  
Sequence 27, Application US/08488351A  
Patent No. 5843446  
GENERAL INFORMATION:  
APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang YI  
APPLICANT: Zamb, Timothy  
TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maria C.H. Lin  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: US  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,351A  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/446,692  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/229,275  
FILING DATE: 14-APR-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/057,166  
FILING DATE: 27-APR-1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria C.H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4146 US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)415-8745  
TELEFAX: (516)751-6849  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-488-351A-27

Query Match 79.6%; Score 134.5; DB 2; Length 33;  
Best Local Similarity 90.3%; Pred. No. 2.7e-13;  
Matches 28; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Oy 2 EKKIAKMEKASSVFNVNNGSPSLHMSYGLRP 32  
Db 3 EKKIAKMEKASSVFNVNNGSGE-HMSYGLRP 32

RESULT 3  
US-08-186-266-6  
Sequence 6, Application US/08186266  
Patent No. 5662907  
GENERAL INFORMATION:  
APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang YI  
APPLICANT: Zamb, Timothy  
TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maria C.H. Lin  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: US  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/186,266  
FILING DATE: 25-JAN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/159,339  
FILING DATE: 29-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,666  
FILING DATE: 07-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 14137-50-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..21  
OTHER INFORMATION: /note= "Plasmodium falciparum CS  
OTHER INFORMATION: protein at positions 378-398."  
US-08-186-266-6

Query Match 52.7%; Score 89; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 9.6e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 EKKIAKMEKASSVFNVNS 20  
Db 3 EKKIAKMEKASSVFNVNS 21

RESULT 4  
US-08-446-692-48  
Sequence 48, Application US/08446692  
Patent No. 5759551  
GENERAL INFORMATION:  
APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang YI  
APPLICANT: Zamb, Timothy  
TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:

TITLE OF INVENTION: INDUCTION OF ANTI-TUMOR CYTOTOXIC  
TITLE OF INVENTION: T LYMPHOCYTES IN HUMANS USING  
TITLE OF INVENTION: SYNTHETIC PEPTIDE EPITOPES  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Hourie and Crew  
STREET: Stuart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/186,266  
FILING DATE: 25-JAN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/159,339  
FILING DATE: 29-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,666  
FILING DATE: 07-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 14137-50-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..21  
OTHER INFORMATION: /note= "Plasmodium falciparum CS  
OTHER INFORMATION: protein at positions 378-398."  
US-08-186-266-6

Query Match 52.7%; Score 89; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 9.6e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 EKKIAKMEKASSVFNVNS 20  
Db 3 EKKIAKMEKASSVFNVNS 21

RESULT 4  
US-08-446-692-48  
Sequence 48, Application US/08446692  
Patent No. 5759551  
GENERAL INFORMATION:  
APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang YI  
APPLICANT: Zamb, Timothy  
TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Maria C.H. Lin  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: US  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,692  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria C.H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4146 US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)415-8745  
TELEFAX: (516)751-6849  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-446-692-48

Query Match 52.7%; Score 89; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 9.6e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNS 20  
|||||  
Db 3 EKKIAKMEKASSVFNVNS 21

RESULT 5  
US-08-488-351A-48  
Sequence 48, Application US/08488351A  
Patent No. 5843446  
GENERAL INFORMATION:  
APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang Yi  
APPLICANT: Zamb, Timothy  
TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maria C.H. Lin  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: US  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,351A  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/446,692  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/229,275  
FILING DATE: 14-APR-1994

CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/057,166  
FILING DATE: 27-APR-1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria C.H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4146 US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)415-8745  
TELEFAX: (516)751-6849  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-488-351A-48

Query Match 52.7%; Score 89; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 9.6e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNS 20  
|||||  
Db 3 EKKIAKMEKASSVFNVNS 21

RESULT 6  
US-09-100-409A-54  
Sequence 54, Application US/09100409A  
Patent No. 6090388  
GENERAL INFORMATION:  
APPLICANT: Wang, Chang Yi  
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR  
TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND  
TITLE OF INVENTION: IMMUNE DISORDERS  
NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentln Release #1.0, Version  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,409A  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME:  
REGISTRATION NUMBER:  
REFERENCE/DOCKET NUMBER: 1151-4154  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-100-409A-54

Query Match 52.7%; Score 89; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 9.6e-07;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIAKMEKASSVFNVNS 20  
|||||

Db 3 EKKIAKMEKASSVFNVNS 21

## RESULT 7

US-08-464-496-17  
Sequence 17, Application US/08464496

Patent No. 6322789

GENERAL INFORMATION:

APPLICANT: EpiImmune, Inc.

APPLICANT: Vitiello, Maria

APPLICANT: Chesnut, Robert

TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL

TITLE OF INVENTION: EPITOPES

FILE REFERENCE: 39963-20001.13

CURRENT APPLICATION NUMBER: US/08/464,496

CURRENT FILING DATE: 1995-06-05

PRIOR APPLICATION NUMBER: 07/935,811

PRIOR FILING DATE: 1992-08-26

PRIOR APPLICATION NUMBER: 07/874,491

PRIOR FILING DATE: 1992-04-27

PRIOR APPLICATION NUMBER: 07/827,682

PRIOR FILING DATE: 1992-01-29

PRIOR APPLICATION NUMBER: 07/749,568

PRIOR FILING DATE: 1991-08-26

NUMBER OF SEQ ID NOS: 75

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17

LENGTH: 21

TYPE: PRT

ORGANISM: Malaria circumsporozoite 378-398

US-08-464-496-17

Query Match 52.7%; Score 89; DB 4; Length 21;

Best Local Similarity 100.0%; Pred. No. 9.6e-07;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIAKMEKASSVFNVNS 20  
|||||

Db 3 EKKIAKMEKASSVFNVNS 21

## RESULT 8

US-08-788-822A-12  
Sequence 12, Application US/08788822A

Patent No. 6413935

GENERAL INFORMATION:

APPLICANT: Alexander, Jeffrey L.

APPLICANT: Defrees, Shawn

APPLICANT: Sette, Alessandro

TITLE OF INVENTION: Induction of Immune Response Against

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/788,822A

FILING DATE: 23-JAN-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/010,510

FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 014137-009210US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-788-822A-12

## Query Match

Best Local Similarity 100.0%; Score 89; DB 4; Length 21;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIAKMEKASSVFNVNS 20  
|||||

Db 3 EKKIAKMEKASSVFNVNS 21

## RESULT 9

US-08-197-484-97

Sequence 97, Application US/08197484

Patent No. 6419931

GENERAL INFORMATION:

APPLICANT: VITIELLO, Maria A.

APPLICANT: CHESNUT, Robert W.

APPLICANT: SETTE, Alessandro D.

APPLICANT: GRAY, Howard

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING

NUMBER OF SEQUENCES: 153

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew

STREET: Steuart Street Tower, One Market Plaza

CITY: San Francisco

STATE: California

COUNTRY: US

ZIP: 94105-1493

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/197,484

FILING DATE: 16-FEB-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/935,811

FILING DATE: 26-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/874,491

FILING DATE: 27-APR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/827,682

FILING DATE: 29-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/749,568

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 14137-26-4

TELECOMMUNICATION INFORMATION:

```
;
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 21 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
;   NAME/KEY: Peptide
;   LOCATION: 1..21
; OTHER INFORMATION: /note="Malaria circumsporozoite
US-08-197-484-97

Query Match      52.7%; Score 89; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.6e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 EKKIAKMEKASSVFNVNS 20
Db      3 EKKIAKMEKASSVFNVNS 21

RESULT 10
PCT-US95-02121-97
; Sequence 97, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 21 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
;   NAME/KEY: Peptide
```

```
;
; LOCATION: 1..21
; OTHER INFORMATION: /note="Malaria circumsporozoite
; OTHER INFORMATION: 378-398"
PCT-US95-02121-97

Query Match      52.7%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.6e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 EKKIAKMEKASSVFNVNS 20
Db      3 EKKIAKMEKASSVFNVNS 21

RESULT 11
PCT-US95-13841-20
; Sequence 20, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic Ige Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 21 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-20

Query Match      52.7%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.6e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 EKKIAKMEKASSVFNVNS 20
Db      3 EKKIAKMEKASSVFNVNS 21

RESULT 12
US-08-760-797A-1
; Sequence 1, Application US/08760797A
; Patent No. 5928902
; GENERAL INFORMATION:
```

APPLICANT: De Wilde, Michel  
APPLICANT: Cohen, Joseph  
TITLE OF INVENTION: Hybrid Protein Between CS  
TITLE OF INVENTION: from Plasmidium and HBSAG  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Smitklime Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/760.797A  
FILING DATE: 04-DEC-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/442.612  
FILING DATE: 17-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: B45015-1C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 423 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-760-797A-1  
Query Match 50.9%; Score 86; DB 2; Length 423;  
Best Local Similarity 85.7%; Pred. NO. 8.1e-05;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 EKKIAXMEKASSVFNVNSGP 22  
Db 174 EKKICKMEKCSSVFNVNSRP 194  
RESULT 13  
US-08-760-797A-3  
Sequence 3, Application US/08760797A  
Patent No. 5928902  
GENERAL INFORMATION:  
APPLICANT: De Wilde, Michel  
APPLICANT: Cohen, Joseph  
TITLE OF INVENTION: Hybrid Protein Between CS  
TITLE OF INVENTION: from Plasmidium and HBSAG  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Smitklime Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/760.797A

FILING DATE: 04-DEC-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/442.612  
FILING DATE: 17-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: B45015-1C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-760-797A-3  
Query Match 50.9%; Score 86; DB 2; Length 424;  
Best Local Similarity 67.7%; Pred. NO. 8.1e-05;  
Matches 21; Conservative 0; Mismatches 4; Indels 6; Gaps 1;  
QY 2 EKKIAXMEKASSVFNVNSGSLHMSYGLRP 32  
Db 171 EKKICKMEKCSSVFNVNS-----SIGLGP 195  
RESULT 14  
US-08-932-929B-1  
Sequence 1, Application US/08932929B  
Patent No. 6169171  
GENERAL INFORMATION:  
APPLICANT: De Wilde, Michel  
APPLICANT: Cohen, Joseph  
TITLE OF INVENTION: Hybrid Protein Between CS  
TITLE OF INVENTION: from Plasmidium and HBSAG  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Smitklime Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/932.929B  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/760.797  
FILING DATE: 04-DEC-1996  
APPLICATION NUMBER: 08/442.612  
FILING DATE: 17-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: B45015-1FWC2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-932-929B-1

Query Match 50.9%; Score 86; DB 3; Length 424;  
Best Local Similarity 85.7%; Pred. No. 8.1e-05;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNNSGP 22  
DB 175 EKKICKMEKSSVFVNNSRP 195

RESULT 15

US-08-932-929B-3  
Sequence 3, Application US/08932929B  
Patent No. 6169171  
GENERAL INFORMATION:  
APPLICANT: De Wilde, Michel  
APPLICANT: Cohen, Joseph  
TITLE OF INVENTION: Hybrid Protein Between CS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/932,929B  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/760,797  
FILING DATE: 04-DEC-1996  
APPLICATION NUMBER: 08/442,612  
FILING DATE: 17-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: B45015-1PMC2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-932-929B-3

Query Match 50.9%; Score 86; DB 3; Length 424;  
Best Local Similarity 67.7%; Pred. No. 8.1e-05;  
Matches 21; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY 2 EKKIAKMEKASSVFVNNSGPSLHWSYGLRP 32  
DB 171 EKKICKMEKSSVFVNNS-----SIGLGP 195

Search completed: September 24, 2003, 17:00:23  
Job time : 10.8812 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:57:10 ; Search time 16.5457 Seconds  
(without alignments)  
301.778 Million cell updates/sec

Title: US-09-848-834A-12

Perfect score: 169

Sequence: 1 DEKKIAKMEKASSVENVVNGPSLHWSYGLRPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb.pap:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb.pap:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pdb.pap:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pdb.pap:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pdb.pap:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pdb.pap:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pdb.pap:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pdb.pap:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pdb.pap:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pdb.pap:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pdb.pap:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pdb.pap:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pdb.pap:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pdb.pap:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pdb.pap:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pdb.pap:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pdb.pap:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pdb.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	168	99.4	33	9	US-09-848-834A-12
2	157	92.9	51	9	US-09-848-834A-20
3	95	56.2	20	9	US-09-848-834A-3
4	95	56.2	36	9	US-09-848-834A-16
5	89	52.7	19	12	US-10-239-313A-54
6	89	52.7	21	12	US-09-932-165-1482
7	89	52.7	21	12	US-10-116-118-33
8	89	52.7	21	12	US-09-935-384-710
9	89	52.7	21	12	US-10-062-109A-761
10	89	52.7	21	12	US-09-942-052-711
11	89	52.7	21	15	US-10-001-469-1404
12	89	52.7	21	15	US-10-128-711-97
13	83.5	49.4	46	9	US-09-848-834A-19
14	82	48.5	21	11	US-09-747-802-23
15	82	48.5	21	11	US-09-865-294-15

16	82	48.5	396	11	US-09-820-843A-31	Sequence 31, Appl
17	81.5	48.2	28	9	US-09-848-834A-11	Sequence 11, Appl
18	79	46.7	17	12	US-10-239-313A-53	Sequence 53, Appl
19	79	46.7	17	15	US-10-128-711-96	Sequence 96, Appl
20	75	44.4	17	12	US-10-239-313A-36	Sequence 36, Appl
21	75	44.4	17	12	US-10-239-313A-38	Sequence 38, Appl
22	75	44.4	17	12	US-10-239-313A-39	Sequence 39, Appl
23	75	44.4	17	12	US-10-239-313A-41	Sequence 41, Appl
24	75	44.4	17	12	US-10-239-313A-60	Sequence 60, Appl
25	75	44.4	17	12	US-10-239-313A-64	Sequence 64, Appl
26	75	44.4	17	12	US-10-239-313A-67	Sequence 67, Appl
27	74	43.8	17	12	US-10-239-313A-59	Sequence 59, Appl
28	74	43.8	17	12	US-10-239-313A-65	Sequence 65, Appl
29	74	43.8	17	12	US-10-239-313A-66	Sequence 66, Appl
30	73	43.2	17	12	US-10-239-313A-14	Sequence 14, Appl
31	73	43.2	17	12	US-10-239-313A-26	Sequence 26, Appl
32	73	43.2	17	12	US-10-239-313A-31	Sequence 31, Appl
33	73	43.2	17	12	US-10-239-313A-32	Sequence 32, Appl
34	73	43.2	17	12	US-10-239-313A-37	Sequence 37, Appl
35	73	43.2	17	12	US-10-239-313A-40	Sequence 40, Appl
36	73	43.2	17	12	US-10-239-313A-44	Sequence 44, Appl
37	73	43.2	17	12	US-10-239-313A-47	Sequence 47, Appl
38	73	43.2	17	12	US-10-239-313A-49	Sequence 49, Appl
39	73	43.2	17	12	US-10-239-313A-52	Sequence 52, Appl
40	73	43.2	17	12	US-10-239-313A-58	Sequence 58, Appl
41	73	43.2	17	12	US-10-239-313A-61	Sequence 61, Appl
42	73	43.2	17	12	US-10-239-313A-69	Sequence 69, Appl
43	73	43.2	17	12	US-10-239-313A-246	Sequence 246, Appl
44	73	43.2	31	9	US-09-848-834A-9	Sequence 9, Appl
45	73	43.2	34	9	US-09-848-834A-10	Sequence 10, Appl

## ALIGNMENTS

### RESULT 1

US-09-848-834A-12  
; Sequence 12, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 33  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 378-398 of  
; OTHER INFORMATION: Plasmodium falciparum CSP protein linked by a spacer to amino  
; OTHER INFORMATION: id sequence 2-10 of the GnRH hormone  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Amidated aspartic acid  
; NAME/KEY: MOD\_RES  
; LOCATION: (33)..(33)  
; OTHER INFORMATION: Amidated glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(20)  
; OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria  
; OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite  
; OTHER INFORMATION: (CSP) protein  
; NAME/KEY: PEPTIDE  
; LOCATION: (21)..(24)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (25)..(33)

```
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-12

Query Match          99.4%; Score 168; DB 9; Length 33;
Best Local Similarity 100.0%; Pred. No. 8.6e-18;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEKKIAKMEKASSVFNVNNGPSLHWSYGLRP 32
   |||||
Db 1 DEKKIAKMEKASSVFNVNNGPSLHWSYGLRP 32

RESULT 2
US-09-848-834A-20
; Sequence 20, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plasmodium
; OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (51)..(51)
; OTHER INFORMATION: Amidated glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(36)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
; OTHER INFORMATION: circumsporozoite (CSP) protein
; NAME/KEY: PEPTIDE
; LOCATION: (37)..(42)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (43)..(51)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

Query Match          92.9%; Score 157; DB 9; Length 51;
Best Local Similarity 94.1%; Pred. No. 5.9e-16;
Matches 32; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

Qy 1 DEKKIAKMEKASSVFNVN--SGPSLHWSYGLRP 32
   |||||
Db 17 DEKKIAKMEKASSVFNVNSSSGPSLHWSYGLRP 50

RESULT 3
US-09-848-834A-3
; Sequence 3, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
```

```
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(20)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium
; OTHER INFORMATION: falciparum circumsporozoite (CSP) protein
US-09-848-834A-3

Query Match          56.2%; Score 95; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEKKIAKMEKASSVFNVNVS 20
   |||||
Db 1 DEKKIAKMEKASSVFNVNVS 20

RESULT 4
US-09-848-834A-16
; Sequence 16, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 378-398
; OTHER INFORMATION: the Plasmodium falciparum circumsporozoite (CSP) protein
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(36)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria
; OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite
; OTHER INFORMATION: (CSP) protein
US-09-848-834A-16

Query Match          56.2%; Score 95; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEKKIAKMEKASSVFNVNVS 20
   |||||
Db 17 DEKKIAKMEKASSVFNVNVS 36
```

RESULT 5  
US-10-239-313A-54  
; Sequence 54, Application US/10239313A  
; Publication NO. US20030175285A1  
; GENERAL INFORMATION:  
; APPLICANT: KLINGUER - HAMOUR, Christine  
; APPLICANT: CORVAIA, Nathalie  
; APPLICANT: BECK, Alain  
; APPLICANT: GOETSCH, Lilliane  
; TITLE OF INVENTION: MOLECULE OF PHARMACEUTICAL INTEREST COMPRISING AT ITS  
; N-TERMINAL A GLUTAMIC ACID OR A GLUTAMINE IN THE FORM  
; OF A PHYSIOLOGICALLY ACCEPTABLE STRONG ACID  
; FILE REFERENCE: 343 727 - US  
; CURRENT APPLICATION NUMBER: US/10/239,313A  
; CURRENT FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: FR 00/03711  
; PRIOR FILING DATE: 2000-03-23  
; PRIOR APPLICATION NUMBER: PCT 01/70772  
; PRIOR FILING DATE: 2001-03-22  
; NUMBER OF SEQ ID NOS: 697  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 54  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Plasmodium malariae  
US-10-239-313A-54

Query Match 52.7%; Score 89; DB 12; Length 19;  
Best Local Similarity 100.0%; Pred. No. 2.2e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKIKAKMEKASSVFNVNS 20  
DB 1 EKIKAKMEKASSVFNVNS 19

RESULT 6  
US-09-932-165-1482  
; Sequence 1482, Application US/09932165  
; Publication NO. US20030134784A1  
; GENERAL INFORMATION:  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: CHALLITA-EID, PIA M.  
; APPLICANT: FARIS, MARY  
; APPLICANT: SAFRAN, DOUGLAS  
; APPLICANT: AFAR, DANIEL  
; APPLICANT: LEVIN, ELANA  
; APPLICANT: HUBERT, RENE  
; APPLICANT: GE, WANGMAO  
; APPLICANT: JAKOBOWITZ, AVA  
; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
; 83P2H3 AND Catif2E11 USEFUL IN TREATMENT AND  
; DETECTION OF CANCER  
; FILE REFERENCE: 51158-20014.00  
; CURRENT APPLICATION NUMBER: US/09/932,165  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/226,329  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 1508  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1482  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-932-165-1482

Query Match 52.7%; Score 89; DB 12; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKIKAKMEKASSVFNVNS 20

DB 3 EKIKAKMEKASSVFNVNS 21  
RESULT 7  
US-10-116-118-33  
; Sequence 33, Application US/10116118  
; Publication NO. US20030143672A1  
; GENERAL INFORMATION:  
; APPLICANT: Tangri, Shabnam  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishicka, Glenn  
; APPLICANT: Fikes, John D.  
; TITLE OF INVENTION: Heteroclitic Analogs and Related Methods  
; FILE REFERENCE: 2060,0090003  
; CURRENT APPLICATION NUMBER: US/10/116,118  
; CURRENT FILING DATE: 2002-08-07  
; PRIOR APPLICATION NUMBER: US 60/166,529  
; PRIOR FILING DATE: 1999-11-18  
; PRIOR APPLICATION NUMBER: US 60/239,008  
; PRIOR FILING DATE: 2000-10-06  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 33  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-10-116-118-33

Query Match 52.7%; Score 89; DB 12; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKIKAKMEKASSVFNVNS 20  
DB 3 EKIKAKMEKASSVFNVNS 21

RESULT 8  
US-09-935-384-710  
; Sequence 710, Application US/09935384  
; Publication NO. US2003016526A1  
; GENERAL INFORMATION:  
; APPLICANT: CHALLITA-EID, PIA  
; APPLICANT: HUBERT, RENE  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: AFAR, DANIEL  
; APPLICANT: LEVIN, ELANA  
; APPLICANT: FARIS, MARY  
; APPLICANT: GE, WANGMAO  
; APPLICANT: JAKOBOWITZ, AVA  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158P1H4  
; USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND  
; OTHER CANCERS  
; FILE REFERENCE: 51158-20033.00  
; CURRENT APPLICATION NUMBER: US/09/935,384  
; CURRENT FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/227,098  
; PRIOR FILING DATE: 2000-08-22  
; NUMBER OF SEQ ID NOS: 783  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 710  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-935-384-710

Query Match 52.7%; Score 89; DB 12; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKIKAKMEKASSVFNVNS 20

Db 3 EKKIAKMEKASSVFNVNS 21

## RESULT 9

US-10-062-109A-761  
; Sequence 761, Application US/10062109A  
; Publication No. US2003016550A1  
; GENERAL INFORMATION:  
; APPLICANT: Ageusys  
; APPLICANT: Challita-Eid, Pia M.  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Morrison, Karen Jane Meyrick  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein  
; TITLE OF INVENTION: Entitled 161P2F10B Useful in Treatment and Detection of  
; TITLE OF INVENTION: Cancer  
; FILE REFERENCE: 51158-20062.01  
; CURRENT APPLICATION NUMBER: US/10/062.109A  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: US 10/005,480  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 765  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 761  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-062-109A-761

Query Match 52.7%; Score 89; DB 12; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNS 20  
|||||  
Db 3 EKKIAKMEKASSVFNVNS 21

## RESULT 10

US-09-942-052-711  
; Sequence 711, Application US/09942052  
; Publication No. US20030170626A1  
; GENERAL INFORMATION:  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Faris, Mary  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Afar, Daniel  
; APPLICANT: Ge, Wangmao  
; APPLICANT: Challita-Eid, Pia M.  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3  
; FILE REFERENCE: 51158-20028.00  
; CURRENT APPLICATION NUMBER: US/09/942.052  
; CURRENT FILING DATE: 2001-08-28  
; PRIOR APPLICATION NUMBER: 60/228,432  
; PRIOR FILING DATE: 2000-08-28  
; NUMBER OF SEQ ID NOS: 744  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 711  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-09-942-052-711

Query Match 52.7%; Score 89; DB 12; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNS 20  
|||||  
Db 3 EKKIAKMEKASSVFNVNS 21

## RESULT 11

US-10-001-469-1404  
; Sequence 1404, Application US/10001469  
; Publication No. US20030091562A1  
; GENERAL INFORMATION:  
; APPLICANT: JAKOBOVITS, AYA  
; APPLICANT: RAITANO, ARTHUR  
; APPLICANT: AFAR, DANIEL  
; APPLICANT: SAFFRAN, DOUGLAS  
; APPLICANT: HUBERT, RENE  
; APPLICANT: FARIS, MARY  
; APPLICANT: CHALLITA-EID, PIA  
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED  
; TITLE OF INVENTION: 101P3A11 USEFUL IN TREATMENT AND DETECTION OF CANCER  
; FILE REFERENCE: 51158-20024.20  
; CURRENT APPLICATION NUMBER: US/10/001.469  
; CURRENT FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: 60/157,902  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/291,118  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: 09/680,728  
; PRIOR FILING DATE: 2000-10-05  
; NUMBER OF SEQ ID NOS: 2888  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1404  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-10-001-469-1404

Query Match 52.7%; Score 89; DB 15; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNS 20  
|||||  
Db 3 EKKIAKMEKASSVFNVNS 21

## RESULT 12

US-10-128-711-97  
; Sequence 97, Application US/10128711  
; Publication No. US20030099634A1  
; GENERAL INFORMATION:  
; APPLICANT: VITIELLO, Maria A.  
; CHESTNUT, Robert W.  
; SETTE, Alessandro D.  
; CELIS, Esteban  
; GRAY, Howard  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/128.711  
; FILING DATE: 22-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/197,484

Sequence	Matches	Mismatches	Indels	Gaps
Sequence 1	18	0	1	0
Sequence 2	18	0	1	0

OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone

QY            2 EKKIAKMEKASSVENVNS 20  
              | | | | | | | | | | | | | | | |  
Db            3 EKKHAKMEKASSVENVNS 21

Search completed: September 24, 2003, 17:33:08  
Job time : 17.5457 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:30:06 ; Search time 11.6094 Seconds  
(without alignments)  
273.361 Million cell updates/sec

Title: US-09-848-834A-12

Perfect score: 169

Sequence: 1 DEKTKAKMEKASSVFNVNSGFLHSYGLRPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	48.5	388	2 A39756	circumsporozoite p
2	82	48.5	405	2 S05428	circumsporozoite p
3	82	48.5	412	1 OZZQAF	circumsporozoite p
4	82	48.5	424	2 A54533	circumsporozoite p
5	82	48.5	442	2 A54529	circumsporozoite p
6	56.5	33.4	92	1 RHRTG	gonadoliberin prec
7	55	32.5	264	2 A44969	circumsporozoite p
8	55	32.5	332	1 OZZOMB	circumsporozoite p
9	55	32.5	348	1 OZZQBK	circumsporozoite p
10	55	32.5	360	2 AD2528	circumsporozoite p
11	55	32.5	367	1 OZZQMY	circumsporozoite p
12	54	32.0	67	2 I78541	gonadoliberin prec
13	54	32.0	90	1 RHMSG	gonadoliberin prec
14	54	32.0	92	1 RHUG	gonadoliberin prec
15	52	30.8	10	1 RHFGG	gonadoliberin - pi
16	52	30.8	10	1 RHSHG	gonadoliberin - pi
17	52	30.8	89	2 IS1423	gonadoliberin prec
18	52	30.8	994	2 S00960	hypothetical prote
19	51	30.2	91	2 JC7393	medaka-type gonad
20	51	30.2	429	2 A54504	circumsporozoite p
21	51	30.2	485	2 A60610	circumsporozoite p
22	51	30.2	507	2 H82378	probable long-chal
23	50	29.6	1484	2 C97196	probable membrane
24	50	29.6	2339	2 A45597	DNA-directed RNA p
25	49.5	29.3	630	2 A36359	VSG expression sit
26	49.5	29.3	630	2 S13724	hypothetical prote
27	49.5	29.3	999	2 S15961	ESAG 8 protein - T
28	49.5	29.3	1265	2 S57968	Ran-binding protei
29	49.5	29.3	3176	1 CGHU3A	collagen alpha 3(V

30 49 29.0 532 2 S27373  
31 49 29.0 1146 2 S40311  
32 48.5 28.7 630 1 BWUT8Q  
33 48 28.4 10 1 RHAQ1  
34 48 28.4 66 2 T43096  
35 48 28.4 92 2 I50644  
36 48 28.4 260 2 T41554  
37 48 28.4 525 2 T40088  
38 48 28.4 3133 2 S52093  
39 47.5 28.1 98 2 I50739  
40 47.5 28.1 562 2 JQ1036  
41 47.5 28.1 664 2 AB1430  
42 47.5 28.1 664 2 AH1803  
43 47.5 28.1 668 1 UYPVME  
44 47.5 28.1 2244 2 F90563  
45 47 27.8 217 2 A98196

beta-fructofuranos  
integrin - fruit f  
regulatory protein  
gonadoliberin I -  
hypothetical prote  
gonadoliberin I pr  
hypothetical prote  
RhogEF domain cont  
hemocytin - silkw  
gonadotropin-relea  
63K adducin-like p  
hypothetical prote  
hypothetical prote  
noncapsid protein  
hypothetical prote  
hypothetical prote

#### ALIGNMENTS

##### RESULT 1

A39756

circumsporozoite protein - Plasmodium reichenowi

C:Species: Plasmodium reichenowi

C:Date: 14-Feb-1992 #sequence\_revision 14-Feb-1992 #text\_change 20-Aug-1999

C:Accession: A39756

R:Lal, A.A.; Goldman, I.F.

J. Biol. Chem. 266, 6686-6689, 1991

A:Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria

A:Reference number: A39756; MUID:91201303; PMID:2016283

A:Accession: A39756

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-388 <LAL>

A:Cross-references: GB:M60972; NID:gl60228; PIDN:AAA29561.1; PID:gl60229

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:312-366/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 48.5%; Score 82; DB 2; Length 388;

Best Local Similarity 78.3%; Pred. NO. 0.00083;

Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKTKAKMEKASSVFNVNSGFLS 24

||||| ||||| ||||| ||||| |||||

Db 356 EKTKAKMEKCSSVFNVNSSIGL 378

##### RESULT 2

S05428

circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)

C:Species: Plasmodium falciparum

C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jun-2000

C:Accession: S05428; A45527; I60657

R:Campbell, J.R.

Nucleic Acids Res. 17, 5854, 1989

A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate

A:Reference number: S05428; MUID:89345189; PMID:2668895

A:Accession: S05428

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-405 <CAM>

A:Cross-references: EMBL:X15363

R:Caspers, P.; Gentz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.

Mol. Biochem. Parasitol. 35, 185-190, 1989

A:Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate

A:Reference number: A45527; MUID:89364998; PMID:2671723

A:Accession: A45527

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-405 <CAS>

A:Cross-references: GB:M22982; GB:J04650; NID:gl60168; PIDN:AAA29527.1; PID:gl60169

R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.



F:24-92/Product: progonadoliberin #status predicted <PQN>  
F:24-33/Product: gonadoliberin #status predicted <GLN>  
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 33.4%; Score 56.5; DB 1; Length 92;  
Best Local Similarity 48.3%; Pred. No. 0.72;  
Matches 14; Conservative 2; Mismatches 12; Indels 1; Gaps 1;

QY 5 IAKMEKASSVFNVNSG--PSLHWSYGLRP 32  
| | | | | | | | | | | | | | | | | | | | | |  
Db 4 IPKLMAAVLLTVCLEGCSSQHSYGLRP 32  
| | | | | | | | | | | | | | | | | | | | | |

RESULT 7  
AD2528  
circumsporozoite protein precursor - Plasmodium yoelii nigeriensis  
C:Species: Plasmodium yoelii nigeriensis  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 01-Dec-1995  
C:Accession: A44969  
R:Colomer-Gould, V.; Enea, V.  
Mol. Biochem. Parasitol. 43, 51-58, 1990  
A:Title: Plasmodium yoelii nigeriensis circumsporozoite gene structure and its implicati  
A:Reference number: A44969; MUID:91148645; PMID:2290446  
A:Accession: A44969  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <COL>  
A:Cross-references: GB:M23250  
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
F:190-242/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 2; Length 264;  
Best Local Similarity 47.4%; Pred. No. 3.9; Mismatches 7; Indels 0; Gaps 0;  
Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNS 20  
: : | | | | | | | | | | | | | | | | | | | | | |  
Db 232 DTEICKMDKCSSIFNIVSN 250  
: : | | | | | | | | | | | | | | | | | | | | | |

RESULT 8  
OZZQMB  
circumsporozoite protein precursor - Plasmodium berghei (strain NK65)  
N:Alternate names: sporozoite surface antigen  
C:Species: Plasmodium berghei  
C:Date: 30-Sep-1987 #sequence\_revision 28-Jul-1995 #text\_change 16-Jul-1999  
C:Accession: A44948; A25083; S13446  
R:Janar, D.E.  
Mol. Biochem. Parasitol. 39, 151-154, 1990  
A:Title: Sequence of the circumsporozoite gene of Plasmodium berghei ANKA clone and NK65  
A:Reference number: A44948; MUID:90158693; PMID:2406593  
A:Accession: A44948  
A:Molecule type: DNA  
A:Residues: 1-332 <LAN>  
A:Cross-references: GB:M28887  
R:Eichinger, D.J.; Arnot, D.E.; Tam, J.P.; Nussenzweig, V.; Enea, V.  
Mol. Cell. Biol. 6, 3965-3972, 1986  
A:Title: Circumsporozoite protein of Plasmodium berghei: gene cloning and identification  
A:Reference number: A25083; MUID:87089740; PMID:2432395  
A:Accession: A25083  
A:Molecule type: DNA  
A:Residues: 1-26, 'I', 28-68, 'PMLRR', 75-126, 'P', 128-134, 'PPPNANDP', 135-332 <ETC>  
A:Cross-references: GB:M14135; MUID:9160245; PIDN:AAA29577.1; PID:gl60246  
R:Weber, J.L.; Egan, J.E.; Lyon, J.A.; Wirtz, R.A.; Charoenvit, Y.; Maloy, W.L.; Hockmey  
Exp. Parasitol. 63, 295-300, 1987  
A:Title: Plasmodium berghei: cloning of the circumsporozoite protein gene.  
A:Reference number: S13446; MUID:87218962; PMID:3556207  
A:Accession: S13446  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 61-122, 'A', 124-332 <WEB>

A:Cross-references: GB:M25445; MUID:gl60177; PIDN:AAA29531.1; PID:gl60178  
C:Comment: There are three distinct regions in the mature circumsporozoite protein, t  
obc membrane-anchoring sequence.  
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
C:Keywords: sporozoite; surface antigen; tandem repeat  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-332/Product: circumsporozoite protein #status predicted <MAT>  
F:94-189/Region: 8-residue repeats  
F:199-230/Region: 2-residue repeats  
F:258-310/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 1; Length 332;  
Best Local Similarity 47.4%; Pred. No. 5;  
Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNS 20  
: : | | | | | | | | | | | | | | | | | | | | | |  
Db 300 DTEICKMDKCSSIFNIVSN 318  
: : | | | | | | | | | | | | | | | | | | | | | |

RESULT 9  
OZZQMB  
circumsporozoite protein precursor - Plasmodium berghei (strain ANKA clone 2.34L)  
N:Alternate names: sporozoite surface antigen  
C:Species: Plasmodium berghei  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Jul-1999  
C:Accession: S07873; S12571  
R:Lockyer, M.J.; Davies, C.S.; Suhrbier, A.; Sinden, R.E.  
Nucleic Acids Res. 18, 376, 1990  
A:Title: Nucleotide sequence of the Plasmodium berghei circumsporozoite protein gene  
A:Reference number: S07873; MUID:90221834; PMID:2183186  
A:Accession: S07873  
A:Molecule type: DNA  
A:Residues: 1-348 <LOC>  
A:Cross-references: EMBL:X17606  
R:Lockyer, M.J.  
Submitted to the EMBL Data Library, November 1989  
A:Reference number: S12571  
A:Accession: S12571  
A:Molecule type: DNA  
A:Residues: 1-59, 'I', 61-81, 83-348 <LOC2>  
A:Cross-references: EMBL:X17606; MUID:99784; PIDN:CAA35608.1; PID:99785  
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
C:Keywords: tandem repeat  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-348/Product: circumsporozoite protein #status predicted <MAT>  
F:94-205/Region: 8-residue repeats  
F:215-247/Region: 2-residue repeats  
F:274-326/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 1; Length 348;  
Best Local Similarity 47.4%; Pred. No. 5.3;  
Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNS 20  
: : | | | | | | | | | | | | | | | | | | | | | |  
Db 316 DTEICKMDKCSSIFNIVSN 334  
: : | | | | | | | | | | | | | | | | | | | | | |

RESULT 10  
AD2528  
Integrase-recombinase protein [imported] - Nostoc sp. (strain PCC 7120) plasmid pCCV1  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AD2528  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irligu  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AD2528  
A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-360 <KUR>  
 A:Cross-references: GB:AP003602; PIDN:BAW77154.1; PID:g17134596; GSPDB:GN00181  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: alr7511  
 A:Genome: plasmid

Query Match 32.5%; Score 55; DB 2; Length 360;  
 Best Local Similarity 47.4%; Pred. No. 5.6;  
 Matches 9; Conservative 7; Mismatches 0; Gaps 0;  
 QY 2 EKKIAKMEKASSVFNVNSGSLHW 26  
 I::IIII::II::: : :  
 Db 164 EREIAKLERAAVADLQHTNKMHV 188

RESULT 11  
 OZZOMY  
 N:Alternate names: protein precursor - Plasmodium yoelii  
 C:Species: Plasmodium yoelii  
 C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 16-Jul-1999  
 C:Accession: A26271  
 R:Lal, A.A.; de la Cruz, V.F.; Welsh, J.A.; Charcoenvit, Y.; Maloy, W.L.; McCutchan, T.F.  
 J. Biol. Chem. 262, 2937-2940, 1987  
 A:Title: Structure of the gene encoding the circumsporozoite protein of Plasmodium yoelii  
 A:Reference number: A26271; MUID:87137555; PMID:3102479  
 A:Accession: A26271  
 A:Molecule type: DNA  
 A:Residues: 1-367 <LAL>  
 A:Cross-references: GB:J02695; NID:g160222; PIDN:AAA29558.1; PID:g160223  
 C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.  
 C:Comment: There are 15 copies of a 6-residue repeat and 8 copies of a 4-residue repeat.  
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 C:Keywords: sporozoite; surface antigen; tandem repeat  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-367/Product: circumsporozoite protein #status predicted <MAT>  
 F:139-228/Region: 6-residue repeats (Q-G-P-G-A-P)  
 F:229-260/Region: 4-residue repeats (Q-Q-P-P)  
 F:293-345/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match 32.5%; Score 55; DB 1; Length 367;  
 Best Local Similarity 47.4%; Pred. No. 5.6;  
 Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNS 20  
 I::II::II::II::II::  
 Db 335 DTEICKMDKCSSIFNIVSN 353

RESULT 12  
 I78541  
 N:Alternate names: rhesus macaque (fragment)  
 A:Title: Luteinizing hormone releasing hormone  
 C:Species: Macaca mulatta (rhesus macaque)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999  
 C:Accession: I78541  
 R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.  
 Neuroendocrinology 60, 346-359, 1994  
 A:Title: Developmental expression of the genes encoding transforming growth factor alpha  
 A:Reference number: I58134; MUID:95124501; PMID:7545971  
 A:Accession: I78541  
 A:Status: preliminary;  
 A:Molecule type: mRNA  
 A:Residues: 1-67 <RES>  
 A:Cross-references: GB:S79518; NID:g912831; PIDN:AAB33096.1; PID:g912832  
 C:Superfamily: gonadoliberin

Query Match 32.0%; Score 54; DB 2; Length 67;  
 Best Local Similarity 90.0%; Pred. No. 1.2;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 SLHWSYGLRP 32  
 I::IIIIIIII  
 Db 5 SQHWSYGLRP 14

RESULT 13  
 RHMSG  
 N:Alternate names: gonadotropin-releasing hormone precursor - mouse  
 N:Contains: gonadoliberin; gonadoliberin-associated protein (GAP)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 31-Dec-1993 #sequence\_revision 18-Mar-1997 #text\_change 18-Jun-1999  
 C:Accession: A47578  
 R:Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Niko  
 Science 234, 1366-1371, 1986  
 A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible  
 A:Reference number: A47578; MUID:87069928; PMID:3024317  
 A:Accession: A47578  
 A:Molecule type: DNA  
 A:Residues: 1-90 <MAS>  
 A:Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175  
 C:Genetics:  
 A:Insertions: 45/3; 77/3  
 C:Function:  
 A:Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropin  
 A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
 C:Superfamily: gonadoliberin  
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:22-31/Product: gonadoliberin #status predicted <GLB>  
 F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>  
 F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
 F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following-

Query Match 32.0%; Score 54; DB 1; Length 90;  
 Best Local Similarity 90.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 SLHWSYGLRP 32  
 I::IIIIIIII  
 Db 21 SQHWSYGLRP 30

RESULT 14  
 RHUG  
 N:Alternate names: gonadoliberin precursor [validated] - human  
 N:Contains: gonadoliberin-releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)  
 C:Species: Homo sapiens (man)  
 C:Date: 17-Mar-1987 #sequence\_revision 21-Jul-1995 #text\_change 08-Dec-2000  
 C:Accession: S05308; A26173; A93342; A90108; A01410; S45718  
 R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.  
 Nucleic Acids Res. 17, 6403-6404, 1989  
 A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone  
 A:Reference number: S05308; MUID:89366682; PMID:2671939  
 A:Accession: S05308  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-92 <HAY>  
 A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956  
 R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
 A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and luteinizing hormone-releasing hormone  
 A:Reference number: A94090; MUID:86094338; PMID:2867548  
 A:Accession: A26173  
 A:Molecule type: mRNA  
 A:Residues: 1-92 <ADE>  
 A:Cross-references: GB:M12578; NID:gi83418; PIDN:AAA35916.1; PID:g386749  
 A:Experimental source: hypothalamus  
 R:Seeburg, P.H.; Adelman, J.P.  
 Nature 311, 666-668, 1984  
 A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasing hormone

A:Reference number: A93342; MUID:85012739; PMID:6090951  
A:Accession: A93342  
A:Molecule type: mRNA  
A:Residues: 1-15, 'S', 17-92 <SEE>  
A:Cross-references: GB:X01059; NID:934356; PIDN:CAA25526.1; PID:934357  
A:Experimental source: placenta  
Rifan, L.; Rousseau, P.  
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982  
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in  
A:Reference number: A90108; MUID:83126573; PMID:6760865  
A:Accession: A90108  
A:Molecule type: protein  
A:Residues: 24-33 <TAN>  
R:Leibowitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterda  
FEBS Lett. 346, 203-206, 1994  
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th  
A:Reference number: S45718; MUID:94283597; PMID:8013634  
A:Contents: annotation; degradation pathway of synthetic hormone  
C:Genetics:  
A:Gene: GDB:GNRH; LHRH; GRH  
A:Cross-references: GDB:I33746; OMIM:227200; OMIM:152760  
A:Map position: 8p21-8p11.2  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: gonadoliberein stimulates pituitary secretion of lutropin and follitropin  
A:Note: gonadoliberein-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberein  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-92/Product: progonaoliberein #status predicted <PGN>  
F:24-33/Product: gonadoliberein #status experimental <MAT>  
F:37-92/Product: gonadoliberein-associated protein #status predicted <GAP>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 32.0%; Score 54; DB 1; Length 92;  
Best Local Similarity 90.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 SLHWSYGLRP 32 -  
| | | | | | | |  
Db 23 SHWSYGLRP 32

RESULT 15  
RHPGG  
gonadoliberein - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 18-Mar-1997  
C:Accession: A01411  
R:Baba, Y.; Matsuo, H.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 44, 459-463, 1971  
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of the  
A:Reference number: A90172; MUID:72114303; PMID:4946067  
A:Accession: A01411  
A:Molecule type: protein  
A:Residues: 1-10 <BAB>  
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 45, 822-827, 1971  
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase metho  
A:Reference number: A90176; MUID:72065376; PMID:4942726  
A:Contents: annotation; synthesis  
A:Note: the synthetic and natural hormones have the same physicochemical and biological  
R:Baba, Y.; Arimura, A.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 45, 483-487, 1971  
A:Title: On the tryptophan residue in porcine LH- and FSH-releasing hormone.  
A:Reference number: A90175; MUID:72117544; PMID:4946275  
A:Contents: annotation  
A:Note: Trp-3 appears to be essential for biological activity  
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo  
C:Superfamily: gonadoliberein  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 30.8%; Score 52; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.26;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32  
| | | | | | | |  
Db 2 HWSYGLRP 9

Search completed: September 24, 2003, 17:49:32  
Job time : 12.6094 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:00:31 ; Search time 6.12465 Seconds  
(without alignments)  
253.383 Million cell updates/sec

Title: US-09-848-834A-12

Perfect score: 169

Sequence: 1 DEKTIKMEKASSVFNVNSGSLHWSYGLRPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	82	48.5	388	1 CSP_PLARE	P26694 plasmodium
2	82	48.5	397	1 CSP_PLAFO	P19597 plasmodium
3	82	48.5	412	1 CSP_PLAFA	P02893 plasmodium
4	82	48.5	424	1 CSP_PLAFT	P13814 plasmodium
5	82	48.5	442	1 CSP_PLAFW	P08307 plasmodium
6	56.5	33.4	92	1 GONL_RAT	P07490 rattus norv
7	55	32.5	339	1 CSP_PLABE	P06915 plasmodium
8	55	32.5	347	1 CSP_PLABA	P23093 plasmodium
9	55	32.5	367	1 CSP_PLAYO	P06914 plasmodium
10	54	32.0	67	1 GONL_MACMU	P55247 macaca mula
11	54	32.0	90	1 GONL_MOUSE	P13562 mus musculu
12	54	32.0	91	1 GONL_PIG	P49921 sus scrofa
13	54	32.0	92	1 GONL_HUMAN	P01148 homo sapien
14	54	32.0	95	1 GONL_PAGMA	P70074 pagrus majo
15	52	30.8	61	1 GONL_SHEEP	Q28588 ovis aries
16	52	30.8	63	1 GONL_MESAU	Q09163 mesocricetu
17	52	30.8	89	1 GONL_XENLA	P45656 xenopus lae
18	52	30.8	92	1 GONL_TUPGB	P45635 tupia glis
19	52	30.8	994	1 DP02_KLUJA	P05468 kluyveromyc
20	51	30.2	92	1 GONL_CAVPO	O54713 cavia porce
21	51	30.2	393	1 CSP_PLABR	P14593 plasmodium
22	51	30.2	429	1 CSP_PLAMA	P13815 plasmodium
23	51	30.2	473	1 SYE_WIGBR	Q84375 wigleswort
24	50	29.6	95	1 GONL_SPAAU	P51919 sparus aura
25	50	29.6	2339	1 RPCI_PLAFA	P27625 plasmodium
26	49.5	29.3	630	1 ESA8_TRYBB	P23799 trypanosoma
27	49.5	29.3	3176	1 CA36_HUMAN	P12111 homo sapien
28	49	29.0	95	1 GONL_MORSA	O73812 morone saxa
29	49	29.0	532	1 INV4_YEAST	P10596 saccharomyc
30	49	29.0	1146	1 ITAL_DROME	Q24247 drosophila
31	48.5	28.7	630	1 ESA8_TRYEQ	P26337 trypanosoma
32	48	28.4	10	1 GONL_ALLMI	P37041 alligator m
33	48	28.4	92	1 GONL_CHICK	P37042 gallus gall

34	48	28.4	99	1 GONL_DICLA	Q91a10 dicentrarch
35	48	28.4	3133	1 HMCT_BOMMO	P98092 bombyx mori
36	47.5	28.1	94	1 GONL_HAPBU	P51918 haplochromi
37	47.5	28.1	668	1 VNCS_MEVA	P27438 mink enteri
38	47.5	28.1	725	1 AADB_MOUSE	Q9qyb8 mus musculu
39	47.5	28.1	725	1 AADB_MOUSE	Q05764 rattus norv
40	47	27.8	275	1 APAR_HAEIN	P44751 haemophilus
41	47	27.8	1032	1 ITA4_XENLA	Q91687 xenopus lae
42	46.5	27.5	110	1 YHJ_ACTAC	P96769 actinobacil
43	46.5	27.5	440	1 FU10_ARATH	Q9sjp6 arabidopsis
44	46.5	27.5	621	1 ULFL_YEAST	Q02724 saccharomyc
45	46.5	27.5	1392	1 LYS2_YEAST	P07702 saccharomyc

## ALIGNMENTS

RESULT 1

ID	CSP_PLARE	STANDARD;	PRT;	388 AA.
AC	P26694;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Circumsporozoite protein precursor (CS).			
OS	Plasmodium reichenowi.			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5854;			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91201303; PubMed=2016283;			
RA	Lal A.A., Goldman I.F.;			
RT	"Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria parasite evolutionarily related to the human malaria parasite Plasmodium falciparum.";			
RL	J. Biol. Chem. 266:6686-6689(1991).			
CC	-1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).			
CC	-1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.			
CC	-1- SIMILARITY: Contains 1 TSP type-1 domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M60972; AAA29561.1; -.			
DR	PIR; A39756; A39756.			
DR	InterPro; IPR003067; Circmsprzoite.			
DR	InterPro; IPR000884; TSP1.			
DR	Pfam; PF00090; TSP_1; 1.			
DR	PRINTS; PR01303; CRCMSPRZOITE.			
DR	SMART; SM00209; TSP1; 1.			
DR	PROSITE; PS50092; TSP1; 1.			
KW	Malaria; Sporozoite; Repeat; Signal.			
FT	SIGNAL 1 16 PROBABLE.			
FT	CHAIN 17 388 CIRCUMSPOROZOITE PROTEIN.			
FT	DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.			
FT	DOMAIN 313 366 TSP TYPE-1.			
SQ	SEQUENCE 388 AA; 42245 MW; C031EEFBE2E35604 CRC64;			

Query Match 48.5%; Score 82; DB 1; Length 388;

Best Local Similarity 78.3%; Pred.No. 0.00034;

Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNVNSGSPSL 24

Db 356 EKKICKMEKSSVFNVNSSIGL 378  
 ||||| ||||| ||||| ||||| |||||

## RESULT 2

CSP\_PLAFO  
 ID CSP\_PLAFO STANDARD; PRT; 397 AA.  
 AC P19597; Q25798;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Circumsporozoite protein precursor (CS).  
 OS Plasmodium falciparum (isolate NF54).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5843;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89345189; PubMed=2668895;  
 RA Campbell J.R.;  
 RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria  
 candidate vaccine antigen.";  
 RL Nucleic Acids Res. 17:5854-5854(1989).  
 RN [2]  
 RP REVISIONS.  
 RA Campbell J.R.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92155298; PubMed=1346766;  
 RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,  
 RA Thomas A.W., Baqar S., Cochran M.A., Thanassi J., Levine M.M.,  
 RA Hackett C.S.;  
 RT "Plasmodium falciparum: in vitro characterization and human  
 infectivity of a cloned line.";  
 RL Exp. Parasitol. 74:159-168(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89364998; PubMed=2671723;  
 RA Caspers P., Gentz R., Matile H., Pink J.R., Sinigaglia F.;  
 RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum  
 isolate used in malaria vaccine trials.";  
 RL Mol. Biochem. Parasitol. 35:185-190(1989).  
 CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT  
 CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE  
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE  
 CC VERTEBRATE HOST).  
 CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR  
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES  
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.  
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X15363; CAA33421.1; -  
 CC EMBL; M83886; AAA29521.1; -  
 CC EMBL; M22982; AAA29527.1; -  
 CC PIR; S05428; S05428.  
 CC InterPro; IPR003067; Crcmsprzoite.  
 CC InterPro; IPR000884; TSP1.  
 CC Pfam; PF00090; tsp\_1; 1.  
 CC PRINTS; PR01303; CRCMSPRZOITE.  
 CC SMART; SM00209; TSP1; 1.  
 CC PROSITE; PS50092; TSP1; 1.  
 CC Malaria; Sporozoite; Repeat; Signal.  
 KW SIGNAL 1 16 PROBABLE.  
 FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.  
 FT DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.

FT DOMAIN 322 375 TSP TYPE-1.  
 FT CONFLICT 194 194 A -> ANPNANPNA (IN REF. 4).  
 SQ SEQUENCE 397 AA; 42646 MW; 9B81146F59BCEA3 CRC64;  
 Query Match 48.5%; Score 82; DB 1; Length 397;  
 Best Local Similarity 78.3%; Pred. No. 0.00034;  
 Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 EKKIAKMEKASSVFNVNSSIGL 24  
 Db 365 EKKICKMEKSSVFNVNSSIGL 387  
 ||||| ||||| ||||| ||||| |||||  
 RESULT 3  
 CSP\_PLAFA  
 ID CSP\_PLAFA STANDARD; PRT; 412 AA.  
 AC P02893;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Circumsporozoite protein precursor (CS).  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84250215; PubMed=6204383;  
 RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,  
 RA Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,  
 RA Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;  
 RT "Structure of the gene encoding the immunodominant surface antigen on  
 the sporozoite of the human malaria parasite Plasmodium falciparum.";  
 RL Science 225:593-599(1984).  
 CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT  
 CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE  
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE  
 CC VERTEBRATE HOST).  
 CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR  
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES  
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.  
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; K02194; AAA29524.1; -  
 CC PIR; A03388; OZZQAF.  
 CC InterPro; IPR003067; Crcmsprzoite.  
 CC InterPro; IPR000884; TSP1.  
 CC Pfam; PF00090; tsp\_1; 1.  
 CC PRINTS; PR01303; CRCMSPRZOITE.  
 CC SMART; SM00209; TSP1; 1.  
 CC PROSITE; PS50092; TSP1; 1.  
 CC Malaria; Sporozoite; Repeat; Signal.  
 KW SIGNAL 1 16 PROBABLE.  
 FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.  
 FT DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.  
 FT DOMAIN 337 390 TSP TYPE-1.  
 SQ SEQUENCE 412 AA; 44420 MW; 1EEED3DE90965F8 CRC64;  
 Query Match 48.5%; Score 82; DB 1; Length 412;  
 Best Local Similarity 78.3%; Pred. No. 0.00036;  
 Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 EKKIAKMEKASSVFNVNSSIGL 24  
 Db 380 EKKICKMEKSSVFNVNSSIGL 402  
 ||||| ||||| ||||| ||||| |||||

```

RESULT 4
CSP_PLAFT STANDARD; PRT; 424 AA.
AC P13614;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS)
OS Plasmodium falciparum (isolate t4 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.";
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19752; AAA29555.1; -
DR PIR; A54533;
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT CHAIN 1 16 PROBABLE.
FT CHAIN 17 424 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT DOMAIN 349 402 TSP TYPE-1.
FT DOMAIN 424 AA; 45610 MW; 710AB14238786CD9 CRC64;
SQ SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;

Query Match 48.5%; Score 82; DB 1; Length 424;
Best Local Similarity 78.3%; Pred. No. 0.00037;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNVSIGL 24
IIII IIII IIIIIIIII I
Db 392 EKKICKMEKCSVFNVNVSIGL 414

RESULT 5
CSP_PLAFT STANDARD; PRT; 442 AA.
AC P08307;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Circumsporozoite protein precursor (CS)
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.";
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19752; AAA29555.1; -
DR PIR; A54533;
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT CHAIN 1 16 PROBABLE.
FT CHAIN 17 424 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT DOMAIN 349 402 TSP TYPE-1.
FT DOMAIN 424 AA; 45610 MW; 710AB14238786CD9 CRC64;
SQ SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;

Query Match 48.5%; Score 82; DB 1; Length 424;
Best Local Similarity 78.3%; Pred. No. 0.00037;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNVSIGL 24
IIII IIII IIIIIIIII I
Db 392 EKKICKMEKCSVFNVNVSIGL 414

RESULT 6
GONI_RAT STANDARD; PRT; 92 AA.
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonaoliberin I precursor [Contains: Gonadoliberin I (LH-RH I) (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor I].
DE GNHR1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat.";
RT factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89384661; PubMed=2476669;

```

```

RX MEDLINE=87115616; PubMed=3543671;
RT Lockyer M.J., Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15505; AAA29554.1; -
DR PIR; A54529; A54529.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT CHAIN 1 16
FT CHAIN 17 442 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 130 320 47 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT DOMAIN 367 420 TSP TYPE-1.
FT DOMAIN 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;
SQ SEQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;

Query Match 48.5%; Score 82; DB 1; Length 442;
Best Local Similarity 78.3%; Pred. No. 0.00039;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNVSIGL 24
IIII IIII IIIIIIIII I
Db 410 EKKICKMEKCSVFNVNVSIGL 432

RESULT 6
GONI_RAT STANDARD; PRT; 92 AA.
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonaoliberin I precursor [Contains: Gonadoliberin I (LH-RH I) (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor I].
DE GNHR1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat.";
RT factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89384661; PubMed=2476669;

```

RA Bond C.T., Haylick J.S., Seeburg P.H., Adelman J.P.;  
 RT "The rat gonadotropin-releasing hormone: SH locus: structure and  
 RL hypothalamic expression.";  
 RN Mol. Endocrinol. 3:1257-1262(1989).  
 RP [3]  
 RC SEQUENCE FROM N.A.  
 RX TISSUE=Thymus;  
 RA MEDLINE=93105480; PubMed=14681115;  
 RT Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;  
 RL "Thymocytes express a mRNA that is identical to hypothalamic  
 RT luteinizing hormone-releasing hormone mRNA.";  
 RN Cell. Mol. Neurobiol. 12:447-454(1992).  
 RP [4]  
 RC SEQUENCE OF 1-47 FROM N.A.  
 RX TISSUE=Heart;  
 RA MEDLINE=87149087; PubMed=3547652;  
 RT Adelman J.P., Bond C.T., Douglass J., Herbert E.;  
 RL "Two mammalian genes transcribed from opposite strands of the same  
 RT DNA locus.";  
 RN Science 235:1514-1517(1987).  
 RP [1]  
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.  
 CC -!- SIMILARITY: Belongs to the GnRH family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; S50870; AAB24572.1; -  
 DR EMBL; M12579; AAA41263.1; -  
 DR EMBL; M31670; AAA41264.1; -  
 DR EMBL; M15527; AAA42141.1; ALT\_SEQ.  
 DR EMBL; M15529; AAA42139.1; -  
 DR EMBL; M15528; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A40147; RHRTG.  
 DR InterPro; IPR002012; GnRH.  
 DR InterPro; IPR004079; GonadoliberinI.  
 DR Pfam; PF00446; GnRH; 1.  
 DR PRINTS; PR01541; GONADOLIBRNI.  
 DR PROSITE; PS00473; GnRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Signal; Pyroglutamate carboxylic acid.  
 FT SIGNAL 1 23  
 FT CHAIN 1 23 PROGONADOLIBERIN I.  
 FT PEPTIDE 24 92 GONADOLIBERIN I.  
 FT PEPTIDE 37 92 PROLACTIN RELEASE-INHIBITING FACTOR I.  
 FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.  
 FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;  
 Query Match 33.4%; Score 56.5; DB 1; Length 92;  
 Best Local Similarity 48.3%; Pred. No. 0.33;  
 Matches 14; Conservative 2; Mismatches 12; Indels 1; Gaps 1;  
 QY 5 IAKMEKASSVFNVNSG-PSLHWSYGLRP 32  
 Db 4 IPKLAAVLLVTCLEGCSSQHSYGLRP 32  
 RESULT 7  
 ID CSP\_PLABE STANDARD; PRT; 339 AA.  
 AC P06915;  
 DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Circumsporozoite protein precursor (CS).  
 OS Plasmodium berghei.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5821;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87089740; PubMed=2432395;  
 RA Eichinger D.J., Arnot D.E., Tam J.P., Nussenzweig V., Enea V.;  
 RT "Circumsporozoite protein of Plasmodium berghei: gene cloning and  
 RT identification of the immunodominant epitopes.";  
 RL Mol. Cell. Biol. 6:3965-3972(1986).  
 CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT  
 CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE  
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE  
 CC VERTEBRATE HOST).  
 CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR  
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES  
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.  
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M14135; AAA29577.1; -  
 DR PIR; A44948; OZQOM.  
 DR InterPro; IPR003067; Crmsprzoite.  
 DR InterPro; IPR000884; TSP1.  
 DR Pfam; PF00090; tsp-1; 1.  
 DR PRINTS; PR01303; CRMSPRZOITE.  
 DR SMART; SM00209; TSP1; 1.  
 DR PROSITE; PS50092; TSP1; 1.  
 KW Malaria; Sporozoite; Repeat; Signal.  
 FT SIGNAL 1 23 PROBABLE.  
 FT CHAIN 24 339 CIRCUMSPOROZOITE PROTEIN.  
 FT DOMAIN 93 196 13 X 8 AA TANDEM REPEATS.  
 FT DOMAIN 206 238 16 X 2 AA TANDEM REPEATS OF P-Q.  
 FT DOMAIN 266 317 TSP TYPE-1.  
 SQ SEQUENCE 339 AA; 37138 MW; E8068A6D11D9551B CRC64;  
 Query Match 32.5%; Score 55; DB 1; Length 339;  
 Best Local Similarity 47.4%; Pred. No. 2.2;  
 Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EKKIAKMEKASSVFNVNS 20  
 Db 307 DTEICKMDKCSIFNIVSN 325  
 RESULT 8  
 ID CSP\_PLABA STANDARD; PRT; 347 AA.  
 AC P23093;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Circumsporozoite protein precursor (CS).  
 OS Plasmodium berghei (strain Anka).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90221834; PubMed=2183186;  
 RA Lockyer M.J., Davies C.S., Suhrbier A., Sinden R.E.;  
 RT "Nucleotide sequence of the plasmodium berghei circumsporozoite  
 RT protein gene from the ANKA clone 2.34L.";  
 RL Nucleic Acids Res. 18:376-376(1990).



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: S75918; AAB33096.1; -  
 DR PIR: I78541; I78541.  
 DR InterPro: IPR002012; GNRH.  
 DR Pfam: PF00446; GNRH: 1.  
 DR PRINTS: PR01541; GONADOLIBRN1.  
 DR PROSITE: PS00473; GNRH: 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal; Pyroglutamate carboxylic acid.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 5 BY SIMILARITY  
 FT CHAIN 6 >67 GONADOLIBERIN I.  
 FT PEPTIDE 6 15 GONADOLIBERIN I.  
 FT ACT\_SITE 19 >67 GNRH-ASSOCIATED PEPTIDE I.  
 FT ACT\_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT MOD\_RES 6 6 ACTIVITY (BY SIMILARITY).  
 FT MOD\_RES 15 15 PYROLIDONE CARBOXYLIC ACID (BY  
 FT NON\_TER 67 67 SIMILARITY).  
 FT SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 32.0%; Score 54; DB 1; Length 67;  
 Best Local Similarity 90.0%; Pred. No. 0.54;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 23 SLHWSYGLRP 32  
 DB 5 SQHWSYGLRP 14  
 I | | | | | | | | | |  
 RESULT 11  
 GONI\_MOUSE  
 ID GONI\_MOUSE STANDARD; PRT; 90 AA.  
 AC P13562;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)  
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor  
 DE I].  
 GN GNRH1 OR GNRH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87069928; PubMed=3024317;  
 RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,  
 RA Phillips H.S., Nikolic K., Seeburg P.H.;  
 RT "A deletion truncating the gonadotropin-releasing hormone gene is  
 RT responsible for hypogonadism in the hpg mouse.";  
 RL Science 234:1366-1371(1986).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC HORMONES.  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the GNRH family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: M14872; AAA37717.1; -  
 DR PIR: A47578; RHMSG.  
 DR MGD; MGI:95789; Gnrh.  
 DR InterPro: IPR002012; GNRH.  
 DR Pfam: PF00446; GNRH: 1.  
 DR PRINTS: PR01541; GONADOLIBRN1.  
 DR PROSITE: PS00473; GNRH: 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Signal; Pyroglutamate carboxylic acid.  
 FT SIGNAL 1 21  
 FT CHAIN 22 90 GONADOLIBERIN I.  
 FT PEPTIDE 22 31 GONADOLIBERIN I.  
 FT ACT\_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACT\_SITE 24 24 ACTIVITY.  
 FT MOD\_RES 22 22 PYROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).  
 FT SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;

Query Match 32.0%; Score 54; DB 1; Length 90;  
 Best Local Similarity 90.0%; Pred. No. 0.74;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 23 SLHWSYGLRP 32  
 DB 21 SQHWSYGLRP 30  
 I | | | | | | | | | |  
 RESULT 12  
 GONI\_PIG  
 ID GONI\_PIG STANDARD; PRT; 91 AA.  
 AC P49921;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)  
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I].  
 GN GNRH1 OR GNRH.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Hypothalamus;  
 RA Weesner G.D., Matteri R.L., Becker B.A.;  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 24-33.  
 RX MEDLINE=72114303; PubMed=4946067;  
 RA Baba Y., Matsuo H., Schally A.V.;  
 RA "Structure of the porcine LH- and FSH-releasing hormone. II.  
 RT Confirmation of the proposed structure by conventional sequential  
 RT analyses.";  
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).  
 RN [3]  
 RP SYNTHESIS OF GONADOLIBERIN.  
 RX MEDLINE=72065376; PubMed=4942726;  
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;  
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-  
 RT phase method.";  
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).  
 RN [4]  
 RP SYNTHESIS OF GONADOLIBERIN.  
 RX MEDLINE=72117544; PubMed=4946275;  
 RA Baba Y., Arimura A., Schally A.V.;  
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";  
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).

CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the GnRH family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L32864; AAA31066.1; -  
 CC InterPro: IPR002012; GnRH.  
 CC InterPro: IPR004079; GonadoliberinI.  
 CC Pfam: PF00446; GnRH; 1.  
 CC PRINTS; PR01541; GONADOLIBRN1.  
 CC PROSITE; PS00473; GnRH; 1.  
 CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Signal; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 23  
 FT CHAIN 24 91 PROGONADOLIBERIN I.  
 FT PEPTIDE 24 33 GONADOLIBERIN I.  
 FT ACT\_SITE 26 26 GnRH-ASSOCIATED PEPTIDE I.  
 FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.  
 FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDA99 CRC64;  
 Query Match 32.0%; Score 54; DB 1; Length 91;  
 Best Local Similarity 90.0%; Pred. No. 0.75;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 23 SLHWSYGLRP 32  
 | | | | | | | | | |  
 Db 23 SQHWSYGLRP 32  
 RESULT 13  
 ID GONL\_HUMAN STANDARD; PRT; 92 AA.  
 AC P01148;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
 DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated  
 DE peptide I].  
 GN GNRH1 OR GnRH OR LHRH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89366682; PubMed=2671939;  
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;  
 RT "The complete nucleotide sequence of the human gonadotropin-releasing  
 RT hormone gene."  
 RL Nucleic Acids Res. 17:6403-6403(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86094338; PubMed=2867548;  
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;  
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor  
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting  
 RT factor in human and rat."  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
 RN [3]

RP SEQUENCE FROM N.A., AND VARIANT SER-16.  
 RX MEDLINE=85012739; PubMed=6090951;  
 RA Seeburg P.H., Adelman J.P.;  
 RT "Characterization of cDNA for precursor of human luteinizing hormone  
 RT releasing hormone."  
 RL Nature 311:666-668(1984).  
 RN [4]  
 RP SEQUENCE OF 24-33.  
 RX MEDLINE=83126573; PubMed=6760865;  
 RA Tan L., Rousseau P.;  
 RT "The chemical identity of the immunoreactive LHRH-like peptide  
 RT biosynthesized in the human placenta."  
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).  
 RN [5]  
 RP VARIANT SER-16.  
 RX MEDLINE=99318093; PubMed=10391209;  
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes."  
 RL Nat. Genet. 22:231-238(1999).  
 RN [6]  
 RP ERRATUM.  
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RL Nat. Genet. 23:373-373(1999).  
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),  
 CC Lutrepulse or Lutrelief (Ferring Pharmaceuticals) and Relisorm  
 CC (Serono).  
 CC -!- SIMILARITY: Belongs to the GnRH family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X01059; CAA25526.1; -  
 CC EMBL; M12578; AAA35916.1; -  
 CC EMBL; X15215; CAA33285.1; -  
 CC PIR: S05308; RHHUG.  
 CC Genew; HGNC:4419; GNRH1.  
 CC MIM; 152760; -  
 CC GO; GO:0003625; C:soluble fraction; TAS.  
 CC GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.  
 CC GO; GO:0007267; P:cell-cell signalling; TAS.  
 CC GO; GO:0007265; P:development; TAS.  
 CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
 CC GO; GO:0007165; P:signal transduction; TAS.  
 CC InterPro: IPR002012; GnRH.  
 CC InterPro: IPR004079; GonadoliberinI.  
 CC Pfam: PF00446; GnRH; 1.  
 CC PRINTS; PR01541; GONADOLIBRN1.  
 CC PROSITE; PS00473; GnRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Pharmaceutical; Signal; Polymorphism;  
 KW Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 23  
 FT CHAIN 24 92 PROGONADOLIBERIN I.  
 FT PEPTIDE 24 33 GONADOLIBERIN I.  
 FT PEPTIDE 37 92 GnRH-ASSOCIATED PEPTIDE I.  
 FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.

FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 FT VARIANT 16 16 W -> S (IN dBSNP:6185).  
 FT /FTid=VAR\_013943.  
 SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;  
 Query Match 32.0%; Score 54; DB 1; Length 92;  
 Best Local Similarity 90.0%; Pred. No. 0.76;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 23 SLHWSYGLRP 32  
 1 | | | | | | | | | |  
 Db 23 SQHWSYGLRP 32  
 1 | | | | | | | | | |  
 RESULT 14  
 GONL\_PAGMA STANDARD; PRT; 95 AA.  
 ID GONL\_PAGMA STANDARD; PRT; 95 AA.  
 AC P70074;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)  
 DE (LH-RH I) (Luliberin I).  
 GN GNRH1.  
 OS Pagrus major (Red sea bream) (Chrysophrys major).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorphia; Perciformes; Percoidae;  
 OC Sparidae; Pagrus.  
 OX NCBI\_TaxID=143350;  
 RN [1]  
 RP TISSUE=Brain;  
 RC Okuzawa K., Granneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;  
 RA Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY  
 SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the GNRH family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; D86582; BAAL3129.1; -  
 DR InterPro; IPR002012; GNRH.  
 DR Pfam; PF00446; GNRH; 1.  
 DR PRINTS; PR01541; GONADOLIBERNI.  
 DR PROSITE; PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Signal; Multigene family; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 95 GONADOLIBERIN I.  
 FT PEPTIDE 24 33 GONADOLIBERIN I.  
 FT PEPTIDE 37 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).  
 FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP)  
 FT (BY SIMILARITY).  
 SQ SEQUENCE 95 AA; 10566 MW; 61E79C99032BD73E CRC64;  
 Query Match 32.0%; Score 54; DB 1; Length 95;  
 Best Local Similarity 52.6%; Pred. No. 0.78;  
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
 OY 14 VFNVNSGSLHWSYGLRP 32  
 1 | : : | | | | | | | |

Db 14 VMVMVSGCCQHWSYGLSP 32  
 RESULT 15  
 GONL\_SHEEP STANDARD; PRT; 61 AA.  
 ID GONL\_SHEEP STANDARD; PRT; 61 AA.  
 AC Q28588;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonadoliberin I precursor (Contains: Gonadoliberin I (LH-RH I)  
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GnRH I) (Luliberin I); GNRH-associated peptide I] (fragment).  
 GN GNRH1 OR GNRH OR LHRH.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE OF 12-61 FROM N.A.  
 RC STRAIN=Western range; TISSUE=Hypothalamus;  
 RA Rodriguez R.E., Wise M.E.;  
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-10.  
 RX MEDLINE=72094314; PubMed=4550508;  
 RA Burgess R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,  
 RA Fellows R., Blackwell R., Vale W., Guillemin R.;  
 RA "Primary structure of the ovine hypothalamic luteinizing hormone-  
 RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass  
 RT spectrometry-decapeptide-Edman degradation).";  
 PL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the GNRH family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U02517; AAA03433.1; -  
 DR InterPro; IPR002012; GNRH.  
 DR Pfam; PF00446; GNRH; 1.  
 DR PRINTS; PR01541; GONADOLIBERNI.  
 DR PROSITE; PS00473; GNRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Pyrrolidone carboxylic acid.  
 FT NON\_TER 1 1  
 FT CHAIN 1 >61 PROGONADOLIBERIN I.  
 FT PEPTIDE 1 10 GONADOLIBERIN I.  
 FT PEPTIDE 14 >61 GNRH-ASSOCIATED PEPTIDE I.  
 FT ACT\_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 ACTIVITY.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).  
 FT NON\_TER 61 61  
 SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;  
 Query Match 30.8%; Score 52; DB 1; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 0.95;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 25 HWSYGLRP 32  
 1 | | : : | | | | | | | |

Db 2 HWSYGLRP 9

Search completed: September 24, 2003, 17:41:18  
Job time : 6.22465 secs

---





```

Q25837
ID Q25837 PRELIMINARY; PRT; 115 AA.
AC Q25837;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=835a;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RT falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83160; AAA29573.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 13018 MW; COA23F5805688237 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00045;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNSIGL 24
Db 83 EKKICKMEKCSSVFNVNNSIGL 105

RESULT 3
Q25839
ID Q25839 PRELIMINARY; PRT; 115 AA.
AC Q25839;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=835c;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RT falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83162; AAA29575.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12974 MW; 6BB538287260DA90 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00045;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNSIGL 24

```

```

Db 83 EKKICKMEKCSSVFNVNNSIGL 105

RESULT 4
Q9U934
ID Q9U934 PRELIMINARY; PRT; 115 AA.
AC Q9U934;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=822a;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RT falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83154; AAA29567.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12925 MW; A24CB33AEBE3F232 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00045;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNSIGL 24
Db 83 EKKICKMEKCSSVFNVNNSIGL 105

RESULT 5
Q25835
ID Q25835 PRELIMINARY; PRT; 115 AA.
AC Q25835;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=828;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RT falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83157; AAA29570.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12925 MW; 9BE52CFEE12BD48F CRC64;

```

```

Query Match      48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00045;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 83 EKKICKMEKCSSVFVNVSIGL 105

RESULT 6
Q25836 PRELIMINARY; PRT; 115 AA.
AC Q25836;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=834b;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RT falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL: M83159; AAA29572.1; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSPI.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSPI.1.
DR PROSITE: PS50092; TSPI.1.
DR NON_TER 1
FT NON_TER 1
SQ SEQUENCE 115 AA; 12926 MW; 697DF630E50B3A90 CRC64;

Query Match      48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00045;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 83 EKKICKMEKCSSVFVNVSIGL 105

RESULT 7
Q9TVW7 PRELIMINARY; PRT; 115 AA.
AC Q9TVW7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=815, 947, and 808;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RT falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL: M83153; AAA29566.1; -.
DR EMBL: M83171; AAA29549.1; -.
DR EMBL: M83151; AAA29564.1; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSPI.

```

```

DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSPI.1.
DR PROSITE: PS50092; TSPI.1.
DR NON_TER 1
SQ SEQUENCE 115 AA; 12925 MW; 697DF63EEBBD90 CRC64;

Query Match      48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00045;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 83 EKKICKMEKCSSVFVNVSIGL 105

RESULT 8
Q25795 PRELIMINARY; PRT; 117 AA.
AC Q25795;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Doolan D.L., Saul A., Good M.F.;
RT "Geographically restricted heterogeneity of the Plasmodium falciparum
RT circumsporozoite protein: relevance for vaccine development.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL: M77203; AAA29517.2; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSPI.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSPI.1.
DR PROSITE: PS50092; TSPI.1.
DR NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 13110 MW; D96BE20944A7C726 CRC64;

Query Match      48.5%; Score 82; DB 5; Length 117;
Best Local Similarity 78.3%; Pred. No. 0.00046;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 85 EKKICKMEKCSSVFVNVSIGL 107

RESULT 9
Q25796 PRELIMINARY; PRT; 117 AA.
AC Q25796;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Doolan D.L., Saul A., Good M.F.;
RT "Geographically restricted heterogeneity of the Plasmodium falciparum
RT circumsporozoite protein: relevance for vaccine development.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL: M77204; AAA29518.2; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSPI.

```

DR InterPro; IPR003067; Crcmsprzoite.

RT  
transmi.  
RL  
Mol. Bio

DR InterPro; IPR003067; Crcmsprzoite.

Query Match 48.5%; Score 82; DB 5; Length 363;  
Best Local Similarity 78.3%; Pred. No. 0.0016;





KW T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;  
 KW scaffold; inhibition; metastasis; wound healing; solid phase.

XX Measles virus.

XX WO9738011-A1.

XX 16-OCT-1997.

XX 03-APR-1997; 97WO-DK00146.

XX 03-APR-1996; 96DK-0000398.

XX (PEPR-) PEPRESEARCH AS.

XX Heegaard PMH, Jakobsen PH;

XX WPI; 1997-512645/47.

PT Non-dendritic peptide carrier linked to a solid phase - useful as a  
 PT diagnostic agent and as a scaffold for production of chemical  
 PT derivatives

XX Claim 30; Page 199; 262pp; English.

CC A non-dendritic peptide carrier (A) has been developed which is coupled  
 CC through a linker to a solid phase, forming a complex of (A)-solid phase.  
 CC Where (A) comprises 10-50 amino acids capable of forming a secondary  
 CC structure in a benign buffer after liberation from the solid phase, and  
 CC further the (A)-solid phase complex comprises an immunogenic substance  
 CC and/or an immune mediator coupled on (A). The present sequence  
 CC represents a specifically claimed T-cell stimulatory peptide from the  
 CC present invention. An (A)-solid phase complex can be used as a scaffold  
 CC for the production of chemical derivatives, characterised by covalently  
 CC attaching molecules at attachment points. Alternatively (A) is used as  
 CC a scaffold-peptide for the incorporation into an immunostimulating  
 CC complex (Iscom) resulting in an (A)-Iscom complex which is used for the  
 CC chemical coupling of antigenic substances in an aqueous solution by  
 CC conjugation. (A) derivatised with one or more peptides having  
 CC fibronectin-, laminin- or vitronectin-like binding activities can be  
 CC used for the promotion of cell-attachment to plastic surfaces, in  
 CC particular to inhibit tumour growth and metastasis, and for promotion  
 CC of wound healing. Also a derivatised (A) can be used for the selection  
 CC of specifically-binding aptamers or as a diagnostic agent. Such  
 CC diagnostic-(A) molecules could be used to detect molecules derived from  
 CC or indicative of pregnancy or of a disease, such as an infectious,  
 CC autoimmune or cancerous disease.  
 CC (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence - 18 AA;

Query Match 44.8%; Score 78; DB 18; Length 18;  
 Best Local Similarity 94.4%; Pred. No. 0.0002;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 17 KLLSEIKGVIVHRLGVE 34  
 Db 1 KLLSLIKGVIVHRLGVE 18  
 |||||  
 |||||

Search completed: September 24, 2003, 17:40:05  
 Job time : 38.4266 secs

```

XX FH Key Location/Qualifiers
XX FT Peptide 1..18
XX FT /note= "Measles virus fusion protein F epitope"
XX FT Peptide 19..22
XX FT /note= "Spacer peptide"
XX FT Peptide 23..31
XX FT /note= "Gonadotrophin releasing hormone epitope"
XX FT Modified-site 31
XX FT /note= "Amidated glycine or glycinamide"
XX PN WC200185763-A2.
XX XX
XX PD 15-NOV-2001.
XX XX
XX PF 04-MAY-2001; 2001WO-0514363.
XX XX
XX PR 05-MAY-2000; 2000US-202328P.
XX XX
XX PA (APHT-) APHTON CORP.
XX XX
XX PI Grimes S, Michaeli D, Stevens VC;
XX XX
XX DR WPI: 2002-049440/06.
XX XX
XX PT Novel synthetic immunogen for inducing immune response against
XX PT gonadotropin releasing hormone, comprises fusion peptide having
XX PT promiscuous helper T-cell peptide epitope and immunomimic peptide
XX PT epitope or its analogue
XX PS Claim 11; Page 7; 43pp; English.
XX XX
XX CC The invention relates to a synthetic immunogen for inducing specific
XX CC antibodies against gonadotropin releasing hormone (GnRH) also known
XX CC as luteinising hormone releasing hormone, LHRH) comprising a fusion
XX CC peptide which comprises a promiscuous helper T-cell peptide epitope and
XX CC immunomimic peptide epitope or its analogue. The synthetic
XX CC immunogen is useful inducing an immune response against GnRH in an
XX CC animal subject, and as such is useful as a contraceptive and in the
XX CC treatment of diseases such as cancer (of the breast, uterus and other
XX CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic
XX CC hypertrophy and prostate cancer. The immunogen is effective in eliciting
XX CC high and specific anti-GnRH antibody titres. The present sequence
XX CC is a synthetic immunogen of the invention.
XX SQ Sequence 31 AA;
XX
XX Query Match 49.4%; Score 86; DB 23; Length 31;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-05;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 17 KLLSEIKGVIVHRLGVE 34
XX |||||
XX Db 1 KLLSEIKGVIVHRLGVE 18
XX
XX RESULT 14
XX ABP72236
XX ID ABP72236 standard; Protein; 75 AA.
XX XX
XX AC ABP72236;
XX XX
XX DT 28-APR-2003 (first entry)
XX XX
XX DE Frizzled putative B-cell epitope-measles virus MVF epitope fusion.
XX XX
XX KW Frizzled; Fzd-2; receptor; human; head and neck squamous carcinoma;
XX KW cancer; diagnosis; immunotherapy; gene therapy; cytostatic;
XX KW antitumour; epitope.
XX XX
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Measles virus.
XX OS Chimeric - Synthetic.

```

```

XX FH Key Location/Qualifiers
XX FT Protein 1..19
XX FT /note= "Measles virus MVF T cell epitope"
XX FT Peptide 20..23
XX FT /note= "Linker peptide"
XX FT Peptide 24..75
XX FT /note= "Frizzled putative B cell epitope"
XX XX
XX PN WC200289081-A2.
XX XX
XX PD 07-NOV-2002.
XX XX
XX PF 01-MAY-2002; 2002WO-0513802.
XX XX
XX PR 01-MAY-2001; 2001US-287995P.
XX XX
XX PA (REGC ) UNIV CALIFORNIA.
XX XX
XX PI Rhee C, Sen M, Wu C, Leoni LM, Corr M, Carson DA;
XX XX
XX DR WPI: 2003-111859/10.
XX DR N-PSDB; AB258288.
XX XX
XX PT Determining overexpression of wnt or frizzled gene in a tumor cell
XX PT useful for immunotherapy in head and neck squamous cell carcinoma
XX PT comprises reverse transcription PCR with primers directed against the
XX PT wnt or frizzled gene
XX PS Example 1; Page 22; 51pp; English.
XX XX
XX CC The present sequence is that of a fusion protein in which the
XX CC measles virus MVF helper T cell epitope is joined via a linker
XX CC peptide to a putative B cell epitope of human frizzled Fzd-2. In an
XX CC example from the invention, DNA encoding the construct was inserted
XX CC into a plasmid vector, and mice were injected with plasmid DNA and
XX CC then boosted with peptide or recombinant protein. The B-cell
XX CC epitope may need to be redesigned to eliminate cross-reactivity in
XX CC the humoral response to other frizzled isoforms. The invention is
XX CC based on the finding that most head and neck squamous carcinoma
XX CC cell lines (HNSCC) overexpress one or more wingless (Wnt) or
XX CC frizzled (Fzd) mRNAs, making the Wnt and Fzd receptors useful
XX CC targets for immunotherapy of this common cancer. The invention
XX CC provides claimed methods for determining, by RT-PCR, overexpression
XX CC of a Wnt or Fzd gene in a tumour cell, for detecting overexpression
XX CC of a Wnt and/or Fzd protein in a cancer, and for altering the
XX CC growth of a cell overexpressing a Wnt and/or Fzd protein by
XX CC contacting the cell with an antibody against the protein, or with a
XX CC synthetic peptide, recombinant protein or DNA vector comprising a
XX CC non-homologous region of Wnt and/or Fzd proteins.
XX SQ Sequence 75 AA;
XX
XX Query Match 46.0%; Score 80; DB 24; Length 75;
XX Best Local Similarity 89.5%; Pred. No. 0.00057;
XX Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 16 LKLLSEIKGVIVHRLGVE 34
XX :|||||
XX Db 1 MKLLSLIKGVIVHRLGVE 19
XX
XX RESULT 15
XX AAW35441
XX ID AAW35441 standard; peptide; 18 AA.
XX XX
XX AC AAW35441;
XX XX
XX DT 25-MAR-2003 (updated)
XX DT 22-APR-1998 (first entry)
XX XX
XX OS T-cell stimulatory peptide from measles virus F-protein 286-302.
XX OS

```

PD 15-NOV-2001.  
 XX  
 PF 04-MAY-2001; 2001WO-US14363.  
 XX  
 PR 05-MAY-2000; 2000US-202328P.  
 XX  
 XX (APHT-) APHTON CORP.  
 XX  
 XX Grimes S, Michaeli D, Stevens VC;  
 PI WPI; 2002-049440/06.  
 DR  
 XX  
 XX Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue  
 XX  
 XX Claim 11; Page 11; 43pp; English.  
 PS  
 XX The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX  
 SQ Sequence 50 AA;  
 Query Match 50.0%; Score 87; DB 23; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 HWSYGLRPGSSGPSL 16  
 DB 2 HWSYGLRPGSSGPSL 16  
 |||||  
 RESULT 12  
 AAU11431  
 ID AAU11431 standard; peptide; 51 AA.  
 XX  
 AC AAU11431;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Synthetic immunogen peptide 12.  
 XX  
 KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Plasmodium falciparum.  
 OS Chimeric - Mammalia.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 1  
 FT /label= OTHER  
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
 FT 1..10  
 FT Peptide  
 FT /note= "Gonadotropin releasing hormone epitope  
 FT (1..10 aa)"  
 FT 11..16  
 FT Peptide  
 FT /note= "Spacer peptide"  
 FT 17..36

FT Peptide /note= "Malaria CSP protein (378-398 aa)"  
 FT 37..42  
 FT /note= "Spacer peptide"  
 FT 43..51  
 FT /note= "Gonadotropin releasing hormone epitope  
 FT (2-10 aa)"  
 FT 51  
 FT /note= "Amidated glycine or glycinamide"  
 XX  
 PN WO200185763-A2.  
 XX  
 XX 15-NOV-2001.  
 PD  
 XX  
 PF 04-MAY-2001; 2001WO-US14363.  
 XX  
 PR 05-MAY-2000; 2000US-202328P.  
 XX  
 XX (APHT-) APHTON CORP.  
 PA  
 XX Grimes S, Michaeli D, Stevens VC;  
 PI WPI; 2002-049440/06.  
 DR  
 XX Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue  
 XX  
 XX Claim 11; Page 12-13; 43pp; English.  
 PS  
 XX The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX  
 SQ Sequence 51 AA;  
 Query Match 50.0%; Score 87; DB 23; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 HWSYGLRPGSSGPSL 16  
 DB 2 HWSYGLRPGSSGPSL 16  
 |||||  
 RESULT 13  
 AAU11420  
 ID AAU11420 standard; peptide; 31 AA.  
 XX  
 AC AAU11420;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Synthetic immunogen peptide 1.  
 XX  
 KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Measles virus.  
 OS Chimeric - Mammalia.  
 OS Synthetic.

PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 XX epitope or its analogue  
 PS Claim 11; Page 10; 43pp; English.  
 XX  
 CC The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX  
 SQ Sequence 36 AA;  
 Query Match 50.0%; Score 87; DB 23; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 HWSYGLRPGSGPSL 16  
 Db 2 HWSYGLRPGSGPSL 16  
 RESULT 10  
 AAU11425  
 ID AAU11425 standard; peptide; 37 AA.  
 AC AAU11425;  
 XX  
 XX 12-MAR-2002 (first entry)  
 DT  
 DE Synthetic immunogen peptide 6.  
 XX  
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Clostridium tetani.  
 OS Chimeric - Mammalia.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 1  
 FT /label= OTHER  
 FT /note= "Pyro-glutamic acid or 5-oxo proline"  
 FT Peptide 1..10  
 FT /note= "Gonadotropin releasing hormone epitope"  
 FT Peptide 11..16  
 FT /note= "Spacer peptide"  
 FT Peptide 17..37  
 FT /note= "Tetanus toxoid sequence (947-967 aa)"  
 XX  
 XX WO200185763-A2.  
 XX  
 XX 15-NOV-2001.  
 PD  
 XX 04-MAY-2001; 2001WO-US14363.  
 PF  
 XX 05-MAY-2000; 2000US-202328P.  
 PR  
 XX (APHT-) APHTON CORP.  
 PA Grimes S, Michaeli D, Stevens VC;  
 XX WPI; 2002-049440/06.  
 XX

XX Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue  
 XX  
 PS Claim 11; Page 9; 43pp; English.  
 XX  
 CC The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX  
 SQ Sequence 37 AA;  
 Query Match 50.0%; Score 87; DB 23; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 HWSYGLRPGSGPSL 16  
 Db 2 HWSYGLRPGSGPSL 16  
 RESULT 11  
 AAU11429  
 ID AAU11429 standard; peptide; 50 AA.  
 AC AAU11429;  
 XX  
 XX 12-MAR-2002 (first entry)  
 DT  
 DE Synthetic immunogen peptide 10.  
 XX  
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Clostridium tetani.  
 OS Chimeric - Mammalia.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 1  
 FT /label= OTHER  
 FT /note= "Pyro-glutamic acid or 5-oxo proline"  
 FT Peptide 1..10  
 FT /note= "Gonadotropin releasing hormone epitope"  
 FT Peptide 11..16  
 FT /note= "Spacer peptide"  
 FT Peptide 17..37  
 FT /note= "Tetanus toxoid (947-967 aa)"  
 XX  
 XX WO200185763-A2.  
 XX  
 XX 15-NOV-2001.  
 PD  
 XX 04-MAY-2001; 2001WO-US14363.  
 PF  
 XX 05-MAY-2000; 2000US-202328P.  
 PR  
 XX (APHT-) APHTON CORP.  
 PA Grimes S, Michaeli D, Stevens VC;  
 XX WPI; 2002-049440/06.  
 XX

CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX

SQ Sequence 31 AA;  
 Query Match 51.1%; Score 89; DB 23; Length 31;  
 Best Local Similarity 83.3%; Pred. No. 9.5e-06;  
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSGPSLKL 19  
 |||||  
 Db 2 HWSYGLRPGSGPSLQYI 19

RESULT 8  
 AAU11430  
 ID AAU11430 standard; peptide; 46 AA.

AC AAU11430;  
 XX  
 DT 12-WAR-2002 (first entry)

XX Synthetic immunogen peptide 11.

KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; GnRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Chimeric - Clostridium tetani.  
 OS Chimeric - Mammalia.  
 OS Synthetic.

FH Key Location/Qualifiers

FT Misc-difference 1  
 FT /label= OTHER  
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
 FT Peptide 1..10 "Gonadotrophin releasing hormone epitope  
 FT (1..10 aa)"

FT Peptide 11..16  
 FT /note= "Spacer peptide"  
 FT Peptide 17..31  
 FT /note= "Tetanus toxoid (830-844 aa)"

FT Peptide 32..37  
 FT /note= "Spacer peptide"  
 FT Peptide 38..46  
 FT /note= "Gonadotrophin releasing hormone epitope  
 FT (2-10 aa)"

FT Modified-site 46  
 FT /note= "Amidated glycine or glycinamide"

FT XX

PN WO200185763-A2.

PD 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US14363.

XX 05-MAY-2000; 2000US-202328P.

XX (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;

XX WPI; 2002-049440/06.

XX Novel synthetic immunogen for inducing immune response against  
 XX gonadotropin releasing hormone, comprises fusion peptide having  
 XX promiscuous helper T-cell peptide epitope and immunomimic peptide  
 XX epitope or its analogue

PS Claim 11; Page 12; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.

SQ Sequence 46 AA;

Query Match 51.1%; Score 89; DB 23; Length 46;  
 Best Local Similarity 83.3%; Pred. No. 1.5e-05;  
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSGPSLKL 19  
 |||||  
 Db 2 HWSYGLRPGSGPSLQYI 19

XX Synthetic immunogen peptide 8.

RESULT 9

AAU11427

ID AAU11427 standard; peptide; 36 AA.

XX AC AAU11427;

XX 12-WAR-2002 (first entry)

XX Synthetic immunogen peptide 8.

XX Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Chimeric - Plasmodium falciparum.

OS Chimeric - Mammalia.

OS Synthetic.

FH Key Location/Qualifiers

FT Misc-difference 1

FT /label= OTHER

FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"

FT Peptide 1..10 "Gonadotrophin releasing hormone epitope"

FT Peptide 11..16 "Spacer peptide"

FT Peptide 17..36 "Malaria CSP protein (378-398 aa)"

FT WO200185763-A2.

PD 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US14363.

XX 05-MAY-2000; 2000US-202328P.

XX (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;

XX WPI; 2002-049440/06.

XX Novel synthetic immunogen for inducing immune response against  
 XX gonadotropin releasing hormone, comprises fusion peptide having

CC putative B cell epitope of human frizzled Fzd-2 is joined via a  
 CC peptide linker to the measles virus MVF T helper epitope. In an  
 CC example from the invention, DNA encoding the construct was inserted  
 CC into a plasmid vector, and mice were injected with plasmid DNA and  
 CC then boosted with peptide or recombinant protein. The B-cell  
 CC epitope may need to be redesigned to eliminate cross-reactivity in  
 CC the humoral response to other frizzled isoforms. The invention is  
 CC based on the finding that most head and neck squamous carcinoma  
 CC cell lines (HNSCC) overexpress one or more wingless (Wnt) or  
 CC frizzled (Fzd) mRNAs, making the Wnt and Fzd receptors useful  
 CC targets for immunotherapy of this common cancer. The invention  
 CC provides claimed methods for determining, by RT-PCR, overexpression  
 CC of a Wnt or Fzd gene in a tumour cell, for detecting overexpression  
 CC of a Wnt and/or Fzd protein in a cancer, and for altering the  
 CC growth of a cell overexpressing a Wnt and/or Fzd protein by  
 CC contacting the cell with an antibody against the protein, or with a  
 CC synthetic peptide, recombinant protein or DNA vector comprising a  
 CC non-homologous region of Wnt and/or Fzd proteins.

XX SQ Sequence 75 AA;

Query Match 58.6%; Score 102; DB 24; Length 75;  
 Best Local Similarity 47.2%; Pred. No. 3.5e-07;  
 Matches 25; Conservative 2; Mismatches 2; Indels 24; Gaps 1;

Qy 6 GLRPGSS-----GPSLKLLSEIKGVIVHRLEGVE 34  
 ||:|||||  
 Db 23 GLQFGAGGTGGPGGGAPPRYATLEHPFHCGPSLKLLSLIKGVIVHRLEGVE 75

RESULT 6  
 AAW67582  
 ID AAW67582 standard; peptide; 40 AA.  
 AC AAW67582;  
 XX  
 XX 02-MAR-1999 (first entry)  
 DE Synthetic chimera fimbriin/T-cell epitope peptide LB2.  
 KW Chimeric; non-typable Haemophilus influenzae; fimbriin; T-cell epitope;  
 KW immunogenic composition; immune response.  
 KW Synthetic.  
 OS  
 XX US5843464-A.  
 PN  
 XX 01-DEC-1998.  
 PD  
 XX 02-JUN-1995; 95US-0460502.  
 PF  
 XX 02-JUN-1995; 95US-0460502.  
 PR  
 XX (OHIS ) UNIV OHIO STATE.  
 PA  
 XX Bakaletz LO, Kaumaya PTP;  
 PI  
 XX WPI; 1999-044514/04.  
 DR  
 XX Synthetic chimeric fimbriin peptide - useful for vaccination against  
 XX non-typable Haemophilus influenzae  
 PT  
 XX Disclosure; Column 4; 16pp; English.  
 PS  
 XX

The invention relates to the manufacture of a synthetic chimeric peptide  
 CC comprising a non-typable Haemophilus influenzae fimbriin peptide fused via  
 CC a linker peptide to a T-cell epitope peptide. The chimeric peptide is  
 CC used in immunogenic compositions which induce an immune response against  
 CC non-typable Haemophilus influenzae. This sequence represents an example  
 CC of a chimeric fimbriin/T-cell epitope peptide and is designated LB2.  
 CC The peptide comprises a 18 amino acid sequence corresponding to amino  
 CC acids 163-180 of the fimbriin protein, the linker sequence and amino acid  
 CC 288-302 of the measles virus fusion protein (a T-cell epitope).

XX SQ Sequence 40 AA;

Query Match 58.0%; Score 101; DB 20; Length 40;  
 Best Local Similarity 87.5%; Pred. No. 2.3e-07;  
 Matches 21; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 11 SSGPSLKLLSEIKGVIVHRLEGVE 34  
 :|||||||  
 Db 17 NTGPSLKLLSLIKGVIVHRLEGVE 40

RESULT 7  
 AAW11426  
 ID AAW11426 standard; peptide; 31 AA.  
 AC AAW11426;  
 XX  
 XX 12-MAR-2002 (first entry)  
 DT  
 XX Synthetic immunogen peptide 7.  
 DE  
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; GnRH; GnRH; GnRH; GnRH; GnRH;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Clostridium tetani.  
 OS Chimeric - Mammalia.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 1  
 FT /label= OTHER  
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
 FT Peptide 1..10  
 FT /note= "Gonadotropin releasing hormone epitope"  
 FT Peptide 11..16  
 FT /note= "Spacer peptide"  
 FT Peptide 17..31  
 FT /note= "Tetanus toxoid sequence (830-844 aa)"  
 XX  
 PN WO200185763-A2.  
 XX  
 XX 15-NOV-2001.  
 PD  
 XX 04-MAY-2001; 2001WO-US14363.  
 PF  
 XX 05-MAY-2000; 2000US-202328P.  
 PR  
 XX (APHT-) APHTON CORP.  
 PA  
 XX Grimes S, Michaeli D, Stevens VC;  
 PI  
 XX WPI; 2002-049440/06.  
 DR  
 XX Novel synthetic immunogen for inducing immune response against  
 XX gonadotropin releasing hormone, comprises fusion peptide having  
 XX promiscuous helper T-cell peptide epitope and immunomimic peptide  
 XX epitope or its analogue  
 PT  
 XX Claim 11; Page 10; 43pp; English.  
 PS  
 XX The invention relates to a synthetic immunogen for inducing specific  
 XX antibodies against gonadotropin releasing hormone (GnRH) also known  
 XX as luteinising hormone releasing hormone (LHRH) comprising a fusion  
 XX peptide which comprises a promiscuous helper T-cell peptide epitope and  
 XX immunomimic peptide epitope or its analogue. The synthetic  
 XX immunogen is useful inducing an immune response against GnRH in an  
 XX animal subject, and as such is useful as a contraceptive and in the  
 XX treatment of diseases such as cancer (of the breast, uterus and other  
 XX gynaecological cancer), endometriosis, uterine fibroids, benign prostatic

XX 02-JUN-1995; 95US-0460502.  
 XX  
 PR 02-JUN-1995; 95US-0460502.  
 XX  
 PA (OHIS ) UNIV OHIO STATE.  
 XX  
 XX Bakaletz LO, Kaumaya PTP;  
 XX  
 DR WPI; 1999-044514/04.  
 XX  
 XX Synthetic chimeric fimbria peptide - useful for vaccination against  
 PT non-typable Haemophilus influenzae  
 XX  
 XX Claim 4; Column 4; 16pp; English.  
 XX  
 CC The invention relates to the manufacture of a synthetic chimeric peptide  
 CC comprising a non-typable Haemophilus influenzae fimbria peptide fused via  
 CC a linker peptide to a T-cell epitope peptide. The chimeric peptide is  
 CC used in immunogenic compositions which induce an immune response against  
 CC non-typable Haemophilus influenzae. This sequence represents an example  
 CC of a chimeric fimbria/T-cell epitope peptide and is designated LBI.  
 CC The peptide comprises a 19 amino acid sequence corresponding to amino  
 CC acids 117-135 of the fimbria protein, the linker sequence and amino acid  
 CC 288-302 of the measles virus fusion protein (a T-cell epitope).  
 XX  
 SQ Sequence 40 AA;  
 Query Match 60.9%; Score 106; DB 20; Length 40;  
 Best Local Similarity 79.3%; Pred. NO. 4.2e-08;  
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 6 GLRPGSSGPKLLSEIKGVIVHRLEGE 34  
 DB 12 GTRDHKKGPKLLSLIKGVIVHRLEGE 40  
 RESULT 4  
 AAY79986  
 ID AAY79986 standard; Peptide; 40 AA.  
 XX  
 AC AAY79986;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE Measles virus fusion protein T-cell promiscuous epitope.  
 XX  
 KW Vaccine; non-typable Haemophilus influenzae; nH1; infection;  
 KW chimeric protein; Haemophilus influenzae; P5-like fimbria protein;  
 KW lipoprotein D; Lbl(f); immunogenic; antigenic; otitis media; sinusitis;  
 KW conjunctivitis; lower respiratory tract infection.  
 XX  
 OS Measles virus.  
 OS Synthetic.  
 XX  
 PN W09964067-A2.  
 XX  
 PD 16-DEC-1999.  
 XX  
 PF 28-MAY-1999; 99WO-US11980.  
 XX  
 PR 11-JUN-1998; 98GB-0012613.  
 XX  
 XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PA (OHIS ) UNIV OHIO STATE RES FOUND.  
 XX  
 XX Bakaletz LO, Cohen J, Dequesne G, Lobet Y;  
 DR WPI; 2000-116457/10.  
 XX  
 XX Novel antigenic P5-like fimbria subunit peptides used in vaccines  
 PT against Haemophilus influenzae  
 XX

PS Example 4; Page 38; 68pp; English.  
 XX  
 CC The present invention describes antigenic P5-like fimbria subunit  
 CC peptides (LBI(f) peptides) of P5-like fimbria proteins from various  
 CC Haemophilus influenzae strains. The peptides are used for diagnosis,  
 CC prevention, and treatment of Haemophilus influenzae infections, such  
 CC as otitis media, sinusitis, conjunctivitis, or lower respiratory tract  
 CC infection. The peptides may also be used in vaccines against  
 CC H. influenzae. Antibodies and probes from the present invention can be  
 CC used for diagnosis of H. influenzae infection. AAY79955 to AAY79993, and  
 CC AA91201 to AA91252, represent sequences used in the exemplification of  
 XX the present invention.  
 SQ Sequence 40 AA;  
 Query Match 60.9%; Score 106; DB 21; Length 40;  
 Best Local Similarity 79.3%; Pred. NO. 4.2e-08;  
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 6 GLRPGSSGPKLLSEIKGVIVHRLEGE 34  
 DB 12 GTRDHKKGPKLLSLIKGVIVHRLEGE 40  
 RESULT 5  
 ABP72235  
 ID ABP72235 standard; Protein; 75 AA.  
 XX  
 AC ABP72235;  
 XX  
 DT 28-APR-2003 (first entry)  
 XX  
 DE Frizzled putative B-cell epitope-measles MVF epitope fusion.  
 XX  
 KW Frizzled; Fzd-2; receptor; human; head and neck squamous carcinoma;  
 KW cancer; diagnosis; immunotherapy; gene therapy; cytostatic;  
 KW antitumour; epitope.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Measles virus.  
 OS Chimeric - Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..53  
 FT /note= "Frizzled putative B cell epitope"  
 FT Peptide 54..57  
 FT /note= "Linker peptide"  
 FT Protein 58..75  
 FT /note= "Measles MVF T cell epitope"  
 PN W0200288081-A2.  
 XX  
 PD 07-NOV-2002.  
 XX  
 PF 01-MAY-2002; 2002WO-US13802.  
 XX  
 PR 01-MAY-2001; 2001US-287995P.  
 XX  
 XX (REGC ) UNIV CALIFORNIA.  
 PA  
 PI Rhee C, Sen M, Wu C, Leoni LM, Corr M, Carson DA;  
 XX  
 DR WPI; 2003-111859/10.  
 DR N-PSDB; ABZ58287.  
 XX  
 XX Determining overexpression of wnt or frizzled gene in a tumor cell  
 PT useful for immunotherapy in head and neck squamous cell carcinoma  
 PT comprises reverse transcription PCR with primers directed against the  
 PT wnt or frizzled gene  
 XX  
 XX Example 1; Page 22; 51pp; English.  
 XX  
 XX The present sequence is that of a fusion protein in which a

```

XX PD 15-NOV-2001.
XX PF
XX PP 04-MAY-2001; 2001WO-US14363.
XX PR
XX PR 05-MAY-2000; 2000US-202328P.
XX PR
XX PA (APHT--) APHTON CORP.
XX PI
XX PI Grimes S, Michaeli D, Stevens VC;
XX DR WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomimic peptide
XX epitope or its analogue
XX
XX Claim 11; Page 9; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific
XX antibodies against gonadotropin releasing hormone (GnRH) also known
XX as luteinising hormone releasing hormone, LHRH) comprising a fusion
XX peptide which comprises a promiscuous helper T-cell peptide epitope and
XX immunomimic peptide epitope or its analogue. The synthetic
XX immunogen is useful inducing an immune response against GnRH in an
XX animal subject, and as such is useful as a contraceptive and in the
XX treatment of diseases such as cancer (of the breast, uterus and other
XX gynaecological cancer), endometriosis, uterine fibroids, benign prostatic
XX hypertrophy and prostate cancer. The immunogen is effective in eliciting
XX high and specific anti-GnRH antibody titres. The present sequence
XX is a synthetic immunogen of the invention.
XX
XX Sequence 34 AA;
XX
XX Query Match 99.4%; Score 173; DB 23; Length 34;
XX Best Local Similarity 100.0%; Pred. No. 5.6e-18;
XX Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 HWSYGLRPGSGPSLKLSEIKGVIVHRLEGVE 34
XX |
XX Db 2 HWSYGLRPGSGPSLKLSEIKGVIVHRLEGVE 34
XX
XX RESULT 2
XX AAU11428
XX ID AAU11428 standard; peptide; 47 AA.
XX AC
XX AC AAU11428;
XX DT
XX DT 12-MAR-2002 (first entry)
XX
XX Synthetic immunogen peptide 9.
XX
XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
XX luteinising hormone releasing hormone; LHRH; contraceptive;
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
XX Chimeric - Plasmodium falciparum.
XX OS
XX OS Chimeric - Mammalia.
XX OS
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 1
XX FT /label= OTHER
XX FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
XX FT
XX FT Peptide 1..10
XX FT /note= "Gonadotropin releasing hormone epitope
XX FT (1..10 aa)"
XX FT Peptide 11..16
XX FT /note= "Spacer peptide"
XX FT

```

```

FT Peptide 17..34
FT /note= "Malaria CSP protein (288-302 aa)"
FT Peptide 35..38
FT /note= "Spacer peptide"
FT Peptide 39..47
FT /note= "Gonadotropin releasing hormone epitope
FT (2-10 aa)"
FT Modified-site 47
FT /note= "Amidated glycine or glycynamide"
XX
XX WO200185763-A2.
XX
XX 15-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US14363.
XX
XX 05-MAY-2000; 2000US-202328P.
XX
XX (APHT--) APHTON CORP.
XX
XX Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomimic peptide
XX epitope or its analogue
XX
XX Claim 11; Page 11; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific
XX antibodies against gonadotropin releasing hormone (GnRH) also known
XX as luteinising hormone releasing hormone, LHRH) comprising a fusion
XX peptide which comprises a promiscuous helper T-cell peptide epitope and
XX immunomimic peptide epitope or its analogue. The synthetic
XX immunogen is useful inducing an immune response against GnRH in an
XX animal subject, and as such is useful as a contraceptive and in the
XX treatment of diseases such as cancer (of the breast, uterus and other
XX gynaecological cancer), endometriosis, uterine fibroids, benign prostatic
XX hypertrophy and prostate cancer. The immunogen is effective in eliciting
XX high and specific anti-GnRH antibody titres. The present sequence
XX is a synthetic immunogen of the invention.
XX
XX Sequence 47 AA;
XX
XX Query Match 99.4%; Score 173; DB 23; Length 47;
XX Best Local Similarity 100.0%; Pred. No. 8.3e-18;
XX Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 HWSYGLRPGSGPSLKLSEIKGVIVHRLEGVE 34
XX |
XX Db 2 HWSYGLRPGSGPSLKLSEIKGVIVHRLEGVE 34
XX
XX RESULT 3
XX AAU67581
XX ID AAU67581 standard; peptide; 40 AA.
XX AC
XX AC AAU67581;
XX DT
XX DT 02-MAR-1999 (first entry)
XX
XX Synthetic chimera fimbria/T-cell epitope peptide LBL.
XX
XX Chimeric; non-typable Haemophilus influenzae; fimbria; T-cell epitope;
XX immunogenic composition; immune response.
XX
XX Synthetic.
XX
XX US5843464-A.
XX
XX 01-DEC-1998.

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:58:36 ; Search time 38.4266 Seconds  
(without alignments)  
140.442 Million cell updates/sec

Title: US-09-848-834A-13

Perfect score: 174

Sequence: 1 XHWSYGLRPGSGPSLKLLSEIKGIVHRLGVGE 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

	1:	2:	3:	4:	5:	6:	7:	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:	24:
	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	173	99.4	34	23 AAU11424	Synthetic Immunoge
2	173	99.4	47	23 AAU11428	Synthetic Immunoge
3	106	60.9	40	20 AAU67581	Synthetic chimera f
4	106	60.9	40	21 AAU79986	Measles virus fusi
5	102	58.6	75	24 ABP72235	Frizzled putative
6	101	58.0	40	20 AAU67582	Synthetic chimera f
7	89	51.1	31	23 AAU11426	Synthetic Immunoge
8	89	51.1	46	23 AAU11430	Synthetic Immunoge
9	87	50.0	36	23 AAU11427	Synthetic Immunoge

10	87	50.0	37	23 AAU11425	Synthetic immunoge
11	87	50.0	50	23 AAU11429	Synthetic immunoge
12	87	50.0	51	23 AAU11431	Synthetic immunoge
13	86	49.4	31	23 AAU11420	Synthetic immunoge
14	80	46.0	75	24 ABP72236	Frizzled putative
15	78	44.8	18	18 AAU35441	T-cell stimulatory
16	77	44.3	16	23 AAU11419	Measles virus fusio
17	77	44.3	25	15 AAU62705	LHRH-containing im
18	74	42.5	20	19 AAU57161	Measles virus mmo
19	74	42.5	20	19 AAU57161	Measles virus F pr
20	73.5	42.2	34	17 AAU05620	miGE2-GG-MVF1Th.
21	73.5	42.2	43	17 AAU05618	miGE1-GG-MVF1Th.
22	73	42.0	42	15 AAU62708	LHRH-containing im
23	72	41.4	15	15 AAU62697	Helper T cell epit
24	72	41.4	15	16 AAU82591	Measles virus F pr
25	72	41.4	15	16 AAU82082	Measles virus glyc
26	72	41.4	15	17 AAU05604	Measles virus F pr
27	72	41.4	15	17 AAU88401	Measles virus F pr
28	72	41.4	15	17 AAU88392	Measles virus F pr
29	72	41.4	15	21 AAU80054	Pathogen derived T
30	72	41.4	15	21 AAU91121	Measles virus F pr
31	72	41.4	15	21 AAU44762	Measles virus prot
32	72	41.4	15	21 AAU68540	Helper T cell epit
33	72	41.4	15	21 AAU54537	T helper cell (Th)
34	72	41.4	15	21 AAU58764	Measles virus F pr
35	72	41.4	15	22 AAB84440	Amino acid sequenc
36	72	41.4	15	22 AAB68638	HER-2 B cell pepti
37	72	41.4	15	23 ABG68169	Pathogen-derived p
38	72	41.4	15	23 ABG68183	Measles virus idea
39	72	41.4	15	23 AAU97873	Measles virus T ce
40	72	41.4	27	15 AAU62707	LHRH-containing im
41	72	41.4	27	16 AAU83584	IgE CH4 region con
42	72	41.4	27	16 AAU83576	IgE CH4 region con
43	72	41.4	27	21 AAU91156	MVF Th epitope/LHR
44	72	41.4	27	21 AAU68567	Peptide immunogen
45	72	41.4	34	17 AAU05619	MVF1Th-GG-miGE2S.

#### ALIGNMENTS

RESULT 1

AAU11424

ID AAU11424 standard; peptide; 34 AA.

XX AAU11424;

AC AAU11424;

XX 12-MAR-2002 (first entry)

DT Synthetic immunogen peptide 5.

DE Gonadotrophin releasing hormone; GnRH; synthetic immunogen;

XX luteinising hormone releasing hormone; LHRH; contraceptive;

KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;

KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;

KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Chimeric - Measles virus.

OS Chimeric - Mammalia.

OS Synthetic.

XX Key

FT Misc-difference 1

FT Location/Qualifiers

FT /label= OTHER

FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"

FT /note= "Gonadotrophin releasing hormone epitope"

FT /note= "Spacer peptide"

FT /note= "Measles virus fusion protein F epitope"

XX WO200185763-A2.

```
Db      1 LSEIKGVIVHRLEGV 15

RESULT 14
US-08-446-692-19
; Sequence 19, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna.
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-19

Query Match      41.4%; Score 72; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19 LSEIKGVIVHRLEGV 33
Db      1 LSEIKGVIVHRLEGV 15

RESULT 15
US-08-488-351A-19
; Sequence 19, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-19

Query Match      41.4%; Score 72; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19 LSEIKGVIVHRLEGV 33
Db      1 LSEIKGVIVHRLEGV 15

Search completed: September 24, 2003, 17:00:23
Job time : 10.1776 secs
```

;; NUMBER OF SEQUENCES: 106  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Morgan & Finnegan, L.L.P.  
;; STREET: 345 Park Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10154-0054  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC Windows  
;; SOFTWARE: Word 97  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/303,323  
;; FILING DATE: 30-APR-1999  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 09/100,414  
;; FILING DATE: 20-JUNE-1998  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Maria H. Lin  
;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4157  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-751-6849  
;; TELEFAX: 212-751-6849  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 15 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-303-323-1

Query Match 41.4%; Score 72; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33  
Db 1 LSEIKGVIVHRLEGV 15

RESULT 12  
US-09-770-014-1  
;; Sequence 1, Application US/09770014  
;; Patent No. 6559282  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, Chang Yi  
;; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
;; TITLE OF INVENTION: IMMUNOGENS  
;; NUMBER OF SEQUENCES: 106  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Morgan & Finnegan, L.L.P.  
;; STREET: 345 Park Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10154-0054  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC Windows  
;; SOFTWARE: Word 97  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/770,014  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 09/100,414  
;; FILING DATE: 20-JUNE-1998  
;; ATTORNEY/AGENT INFORMATION:

;; NAME: Maria H. Lin  
;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4157  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-751-6800  
;; TELEFAX: 212-751-6849  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 15 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-770-014-1

Query Match 41.4%; Score 72; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33  
Db 1 LSEIKGVIVHRLEGV 15

RESULT 13  
PCT-US95-13841-12  
;; Sequence 12, Application PC/TUS9513841  
;; GENERAL INFORMATION:  
;; APPLICANT: United Biomedical Inc; Walfield, Alan M.;  
;; APPLICANT: Wang, Chang Yi  
;; TITLE OF INVENTION: Synthetic IgE Membrane Anchor  
;; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy  
;; NUMBER OF SEQUENCES: 32  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Maria C.H. Lin  
;; STREET: 345 Park Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10154  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: WordPerfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/13841  
;; FILING DATE: 25-OCT-1995  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/328,519  
;; FILING DATE: 25-OCT-1994  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Lin, Maria C.H.  
;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4117  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-751-6800  
;; TELEFAX: 212-751-6849  
;; TELEX: 421792  
;; INFORMATION FOR SEQ ID NO: 12:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 15 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
PCT-US95-13841-12

Query Match 41.4%; Score 72; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33

; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/446,692  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/229,275  
; FILING DATE: 14-APR-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/057,166  
; FILING DATE: 27-APR-1992  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria C.H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)415-8745  
; TELEFAX: (516)751-6849  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-488-351A-20

Query Match 42.08; Score 73; DB 2; Length 42;  
Best Local Similarity 93.88; Pred. No. 0.00066;  
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 18 LLSEIKGVIVHRLEGV 33  
:|||||  
Db 15 VLSEIKGVIVHRLEGV 30

RESULT 9  
US-09-100-414B-1  
; Sequence 1, Application US/09100414B  
; Patent No. 6025468  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi  
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
; TITLE OF INVENTION: IMMUNOGENS  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Finnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10154-0054  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC Windows  
; SOFTWARE: Word 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/100,414B  
; FILING DATE: 20-JUNE-1998  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4157  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-100-414B-1

Query Match 41.48; Score 72; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33  
:|||||  
Db 1 LSEIKGVIVHRLEGV 15

RESULT 10  
US-09-100-409A-38  
; Sequence 38, Application US/09100409A  
; Patent No. 6090388  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi  
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR  
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND  
; TITLE OF INVENTION: IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10154-0054  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version  
; SOFTWARE: #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/100,409A  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME:  
; REGISTRATION NUMBER:  
; REFERENCE/DOCKET NUMBER: 1151-4154  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-100-409A-38

Query Match 41.48; Score 72; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33  
:|||||  
Db 1 LSEIKGVIVHRLEGV 15

RESULT 11  
US-09-303-323-1  
; Sequence 1, Application US/09303323  
; Patent No. 6228987  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi  
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
; TITLE OF INVENTION: IMMUNOGENS

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-28

Query Match 42.2%; Score 73.5; DB 5; Length 34;
Best Local Similarity 64.3%; Pred. No. 0.00043;
Matches 18; Conservative 1; Mismatches 4; Indels 5; Gaps 1;

QY 6 GLRPGSSGSLKLLSEIKGVIVHRLEGV 33
   | | | | | | | | | | | | | | | |
Db 12 GEAPWTGG-----LSEIKGVIVHRLEGV 34

RESULT 6
PCT-US95-13841-26
; Sequence 26, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-26

Query Match 42.2%; Score 73.5; DB 5; Length 43;
Best Local Similarity 64.3%; Pred. No. 0.00057;
Matches 18; Conservative 1; Mismatches 4; Indels 5; Gaps 1;

QY 6 GLRPGSSGSLKLLSEIKGVIVHRLEGV 33
   | | | | | | | | | | | | | | | |
Db 21 GQQQQLGG-----LSEIKGVIVHRLEGV 43

RESULT 7
US-08-446-692-20
; Sequence 20, Application US/08446692
```

```
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-20

Query Match 42.0%; Score 73; DB 1; Length 42;
Best Local Similarity 93.8%; Pred. No. 0.00066;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 18 LLSEIKGVIVHRLEGV 33
   | | | | | | | | | | | | | | | |
Db 15 VLSEIKGVIVHRLEGV 30

RESULT 8
US-08-488-351A-20
; Sequence 20, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
```

PCT/US95-13841-28  
 ; Sequence 28, Application PC/TUS9513841  
 ; GENERAL INFORMATION:  
 ; APPLICANT: United Biomedical Inc; Walfield, Alan M.;  
 ; APPLICANT: Wang, Chang Yi  
 ; TITLE OF INVENTION: Synthetic IgE Membrane Anchor  
 ; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy  
 ; NUMBER OF SEQUENCES: 32  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Maria C.H. Lin  
 ; STREET: 345 Park Avenue  
 ; CITY: New York  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: WordPerfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/13841  
 ; FILING DATE: 25-OCT-1995  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/328,519  
 ; FILING DATE: 25-OCT-1994  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lin, Maria C.H.  
 ; REGISTRATION NUMBER: 29 323  
 ; REFERENCE/DOCKET NUMBER: 1151-4117  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-758-4800  
 ; TELEFAX: 212-751-6849  
 ; TELEX: 421792  
 ; INFORMATION FOR SEQ ID NO: 28:

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:46:20 ; Search time 10.0776 Seconds  
(without alignments)  
142.750 Million cell updates/sec

Title: US-09-848-834A-13  
Perfect score: 174  
Sequence: 1 XHWSYGLRPGSSGSLKLLSEIKGVIVHRLGVE 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues  
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PT05\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	60.9	40	2	US-08-460-502-10
2	106	60.9	40	4	US-09-148-711A-10
3	101	58.0	40	2	US-08-460-502-11
4	96	55.2	40	4	US-09-148-711A-11
5	73.5	42.2	34	5	PCT-US95-13841-28
6	73.5	42.2	43	5	PCT-US95-13841-26
7	73	42.0	42	1	US-08-446-692-20
8	73	42.0	42	2	US-08-488-351A-20
9	72	41.4	15	3	US-09-100-414B-1
10	72	41.4	15	3	US-09-100-409A-38
11	72	41.4	15	3	US-09-303-323-1
12	72	41.4	15	4	US-09-770-014-1
13	72	41.4	15	5	PCT-US95-13841-12
14	72	41.4	27	1	US-08-446-692-19
15	72	41.4	27	2	US-08-488-351A-19
16	72	41.4	27	3	US-09-100-414B-36
17	72	41.4	27	3	US-09-303-323-36
18	72	41.4	27	4	US-09-770-014-36
19	72	41.4	34	5	PCT-US95-13841-27
20	72	41.4	35	1	US-08-446-692-55
21	72	41.4	35	1	US-08-446-692-61
22	72	41.4	35	2	US-08-488-351A-55
23	72	41.4	35	2	US-08-488-351A-61
24	72	41.4	36	3	US-09-082-279B-505
25	72	41.4	36	4	US-09-315-304B-505
26	72	41.4	36	4	US-09-834-784-505
27	72	41.4	42	5	PCT-US95-13841-25

28	72	41.4	45	1	US-08-446-692-33	Sequence 33, Appl
29	72	41.4	45	2	US-08-488-351A-33	Sequence 33, Appl
30	72	41.4	438	3	US-08-486-099-105	Sequence 105, App
31	72	41.4	438	3	US-08-360-107A-115	Sequence 115, App
32	72	41.4	438	3	US-08-484-223B-105	Sequence 105, App
33	72	41.4	438	3	US-08-919-597-105	Sequence 105, App
34	72	41.4	438	3	US-08-475-668A-105	Sequence 105, App
35	72	41.4	438	3	US-08-485-551A-105	Sequence 105, App
36	72	41.4	438	3	US-08-471-913A-105	Sequence 105, App
37	72	41.4	438	3	US-08-485-264A-105	Sequence 105, App
38	72	41.4	438	4	US-08-474-349A-105	Sequence 105, App
39	72	41.4	438	4	US-08-470-896-105	Sequence 105, App
40	72	41.4	438	4	US-08-485-546A-105	Sequence 105, App
41	72	41.4	550	1	US-08-279-700-16	Sequence 16, Appl
42	72	41.4	550	1	US-08-279-700-18	Sequence 18, Appl
43	72	41.4	550	1	US-08-279-700-20	Sequence 20, Appl
44	72	41.4	550	1	US-08-279-700-22	Sequence 22, Appl
45	72	41.4	550	1	US-08-348-891A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1  
US-08-460-502-10  
; Sequence 10, Application US/08460502  
; Patent No. 5843464  
; GENERAL INFORMATION:  
; APPLICANT: Bakaletz, Lauren O.  
; APPLICANT: Kaumaya, Parvin T.  
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Calfee, Halter and Griswold  
; CITY: Cleveland  
; STATE: Ohio  
; COUNTRY: U.S.A.  
; ZIP: 44114-2688  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,502  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goltick, Mary E.  
; REGISTRATION NUMBER: 34,829  
; REFERENCE/DOCKET NUMBER: 22727/00120  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (216) 622-8458  
; TELEFAX: (216) 241-0816  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-460-502-10

Query Match 60.9%; Score 106; DB 2; Length 40;  
Best Local Similarity 79.3%; Pred. No. 7.9e-09;  
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 6 GLRPGSSGSLKLLSEIKGVIVHRLGVE 34  
DB 12 GTRDHHKGPGLKLLSLKGVIVHRLGVE 40

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PMVF-FZD2
US-09-847-102A-33

Query Match      46.0%; Score 80; DB 11; Length 75;
Best Local Similarity 89.5%; Pred. No. 0.00077;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 16 LKLLSEIKGVIVHRLEGVE 34
Db 1 MKLLSLIKGVIVHRLEGVE 19

RESULT 15
US-10-285-976-231
; Sequence 231, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Malini, Sen
; APPLICANT: Wu, Christina
; APPLICANT: Leoni, Lorenzo M.
; APPLICANT: Corr, Maripat
; APPLICANT: Carson, Dennis A.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
; TITLE OF INVENTION: In Head and Neck Squamous Cell Carcinomas
; FILE REFERENCE: 023070-130320US
; CURRENT APPLICATION NUMBER: US/10/285,976
; CURRENT FILING DATE: 2002-11-01
; PRIOR FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: US 60/287,995
; PRIOR FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: WO PCT/US02/13802
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 231
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:PMVF-ZD2
; OTHER INFORMATION: measles virus fusion (MVF) epitope fused to
; OTHER INFORMATION: frizzled domain
US-10-285-976-231

Query Match      46.0%; Score 80; DB 12; Length 75;
Best Local Similarity 89.5%; Pred. No. 0.00077;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 16 LKLLSEIKGVIVHRLEGVE 34
Db 1 MKLLSLIKGVIVHRLEGVE 19

Search completed: September 24, 2003, 17:33:08
Job time : 17.0471 secs

```

```

; NAME/KEY: PEPTIDE
; LOCATION: (17)..(37)
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent
; OTHER INFORMATION: oxylysine
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(41)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (42)..(50)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

Query Match          50.0%; Score 87; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HWSYGLRPGSGPSL 16
        |||||
DB      2 HWSYGLRPGSGPSL 16

RESULT 12
US-09-848-834A-20
; Sequence 20, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 20
; TYPE: PRT
; LENGTH: 51
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plasmod
; OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (51)..(51)
; OTHER INFORMATION: Amidated glycine or glycylamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(36)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
; NAME/KEY: PEPTIDE
; LOCATION: (37)..(42)
; OTHER INFORMATION: circumsporozoite (CSP) protein
; NAME/KEY: PEPTIDE
; LOCATION: (43)..(51)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

Query Match          50.0%; Score 87; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 5.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HWSYGLRPGSGPSL 16
        |||||
DB      2 HWSYGLRPGSGPSL 16

RESULT 13
US-09-848-834A-9
; Sequence 9, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 9
; TYPE: PRT
; LENGTH: 31
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide made up of amino acid sequence 288-302 of the
; OTHER INFORMATION: asels virus fusion protein, F linked by a spacer peptide to an
; OTHER INFORMATION: acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated Lysine
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(15)
; OTHER INFORMATION: Peptide corresponds to the amino acid sequences 288-302 of the
; OTHER INFORMATION: measles virus fusion protein, F
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(22)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (23)..(31)
; OTHER INFORMATION: Peptide corresponds to amino acid sequences 2-10 of the human
; OTHER INFORMATION: GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (31)..(31)
; OTHER INFORMATION: Amidated glycine or glycylamide
US-09-848-834A-9

Query Match          49.4%; Score 86; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17 KLLSEIKGVIVHRLEGVE 34
        |||||
DB      1 KLLSEIKGVIVHRLEGVE 18

RESULT 14
US-09-847-102A-33
; Sequence 33, Application US/09847102A
; Publication No. US2003004409A1
; GENERAL INFORMATION:
; APPLICANT: University of California
; APPLICANT: Carson, Dennis A.
; APPLICANT: Corr, Maripat
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Lorenzo, Leonl M.
; APPLICANT: Malini, Sen
; TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 22000-20629.00
; CURRENT APPLICATION NUMBER: US/09/847,102A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 75
```

```

Db      2 HWSYGLRPGSGPSL 16

RESULT 13
US-09-848-834A-9
; Sequence 9, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 9
; TYPE: PRT
; LENGTH: 31
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide made up of amino acid sequence 288-302 of the
; OTHER INFORMATION: asels virus fusion protein, F linked by a spacer peptide to an
; OTHER INFORMATION: acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated Lysine
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(15)
; OTHER INFORMATION: Peptide corresponds to the amino acid sequences 288-302 of the
; OTHER INFORMATION: measles virus fusion protein, F
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(22)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (23)..(31)
; OTHER INFORMATION: Peptide corresponds to amino acid sequences 2-10 of the human
; OTHER INFORMATION: GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (31)..(31)
; OTHER INFORMATION: Amidated glycine or glycylamide
US-09-848-834A-9

Query Match          49.4%; Score 86; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17 KLLSEIKGVIVHRLEGVE 34
        |||||
DB      1 KLLSEIKGVIVHRLEGVE 18

RESULT 14
US-09-847-102A-33
; Sequence 33, Application US/09847102A
; Publication No. US2003004409A1
; GENERAL INFORMATION:
; APPLICANT: University of California
; APPLICANT: Carson, Dennis A.
; APPLICANT: Corr, Maripat
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Lorenzo, Leonl M.
; APPLICANT: Malini, Sen
; TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 22000-20629.00
; CURRENT APPLICATION NUMBER: US/09/847,102A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 75
```



```
; FILE REFERENCE: 023070-130320US
; CURRENT APPLICATION NUMBER: US/10/285,976
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US 60/287,995
; PRIOR FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: WO PCT/US02/13802
; PRIOR FILING DATE: 2002-05-01
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 229
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:PF2D2-MWVF
; OTHER INFORMATION: measles virus fusion (MWV) epitope fused to
; OTHER INFORMATION: frizzled domain
US-10-285-976-229

Query Match          58.6%; Score 102; DB 12; Length 75;
Best Local Similarity 47.2%; Pred No. 6.2e-07;
Matches 25; Conservative 2; Mismatches 2; Indels 24; Gaps 1;

QY 6 GLRPGSS-----GPSLKLSEIKGVIVHRLEGVE 34
|||:|
Db 23 GLQPGAGGTPGGPGGGAPPRVATLEHPHCGPSLKLSEIKGVIVHRLEGVE 75

RESULT 6
US-10-223-711-11
; Sequence 11, Application US/10223711
; Publication No. US20030113344A1
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Pravin T.P.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; FILE REFERENCE: 18525/04058
; CURRENT APPLICATION NUMBER: US/10/223,711
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 09/148,711
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 08/460,502
; PRIOR FILING DATE: 1995-06-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-223-711-11

Query Match          55.2%; Score 96; DB 15; Length 40;
Best Local Similarity 87.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 SSGPSLKLSEIKGVIVHRLEGV 33
:::||||| |||||
Db 17 NTGPSLKLSEIKGVIVHRLEGV 39

RESULT 7
US-09-848-834A-15
; Sequence 15, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetanus
; OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino acid
; OTHER INFORMATION: sequence 1-10 of GnRH
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (46)..(46)
; OTHER INFORMATION: Amidated glycine or glycynamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
```

```
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 830-844 of Tetanus
; OTHER INFORMATION: tanus toxoid precursor (Tentoxylisin)
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
US-09-848-834A-15

Query Match          51.1%; Score 89; DB 9; Length 31;
Best Local Similarity 83.3%; Pred. No. 1.5e-05;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSGPSLKL 19
|||||
Db 2 HWSYGLRPGSSGPSLOYI 19

RESULT 8
US-09-848-834A-19
; Sequence 19, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetanus
; OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino acid
; OTHER INFORMATION: sequence 1-10 of GnRH
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (46)..(46)
; OTHER INFORMATION: Amidated glycine or glycynamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
```

```
Best Local Similarity 100.0%; Pred. No. 2.6e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSGLRPGSGPSLKLSEIKGVIVHRLGVE 34
    |||||
Db 2 HWSGLRPGSGPSLKLSEIKGVIVHRLGVE 34

RESULT 2
US-09-848-834A-17
; Sequence 17, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848.834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the G
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
; OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid seq
; OTHER INFORMATION: uence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (47)..(47)
; OTHER INFORMATION: Amidated-glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(18)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(34)
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,
; NAME/KEY: PEPTIDE
; LOCATION: (35)..(38)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (39)..(47)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17

Query Match 99.4%; Score 173; DB 9; Length 47;
Best Local Similarity 100.0%; Pred. No. 3.7e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSGLRPGSGPSLKLSEIKGVIVHRLGVE 34
    |||||
Db 2 HWSGLRPGSGPSLKLSEIKGVIVHRLGVE 34

RESULT 3
US-10-223-711-10
; Sequence 10, Application US/10223711
; Publication No. US20030113344A1
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Pravin T.P.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; FILE REFERENCE: 18525/04058
; CURRENT APPLICATION NUMBER: US/10/223,711

; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 09/148,711
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 08/460,502
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic construct
US-10-223-711-10

Query Match 60.9%; Score 106; DB 15; Length 40;
Best Local Similarity 79.3%; Pred. No. 8.2e-08;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 6 GLRPGSGPSLKLSEIKGVIVHRLGVE 34
    |||||
Db 12 GTRDHKKGPSLKLSEIKGVIVHRLGVE 40

RESULT 4
US-09-847-102A-31
; Sequence 31, Application US/09847102A
; Publication No. US20030044409A1
; GENERAL INFORMATION:
; APPLICANT: University of California
; APPLICANT: Carson, Dennis A.
; APPLICANT: Corr, Maripat
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Lorenzo, Leoni M.
; APPLICANT: Malini, Sen
; TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS
; FILE REFERENCE: 22000-20629.00
; CURRENT APPLICATION NUMBER: US/09/847.102A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: PFZD2-MMVF
US-09-847-102A-31

Query Match 58.6%; Score 102; DB 11; Length 75;
Best Local Similarity 47.2%; Pred. No. 6.2e-07;
Matches 25; Conservative 2; Mismatches 2; Indels 24; Gaps 1;

Qy 6 GLRPGSS-----GPSLKLSEIKGVIVHRLGVE 34
    |||||
Db 23 GLOPGAGGTGGPGGGAPPRYATLEHPFHCGPSLKLSEIKGVIVHRLGVE 75

RESULT 5
US-10-285-976-229
; Sequence 229, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Malini, Sen
; APPLICANT: Wu, Christina
; APPLICANT: Leoni, Lorenzo M.
; APPLICANT: Corr, Maripat
; APPLICANT: Carson, Dennis A.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
; TITLE OF INVENTION: in Head and Neck Squamous Cell Carcinomas
```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:57:10 ; Search time 17.0471 Seconds  
(without alignments)  
301.778 Million cell updates/sec

Title: US-09-848-834A-13

Perfect score: 174

Sequence: 1 XHWSYGLRPGSSGSLKLLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PTC\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09C\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10C\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	173	99.4	34	9 US-09-848-834A-13	Sequence 13, Appl
2	173	99.4	47	9 US-09-848-834A-17	Sequence 17, Appl
3	106	60.9	40	15 US-10-223-711-10	Sequence 10, Appl
4	102	58.6	75	11 US-09-847-102A-31	Sequence 31, Appl
5	102	58.6	75	12 US-10-285-976-229	Sequence 229, Appl
6	96	55.2	40	15 US-10-223-711-11	Sequence 11, Appl
7	89	51.1	31	9 US-09-848-834A-15	Sequence 15, Appl
8	89	51.1	46	9 US-09-848-834A-19	Sequence 19, Appl
9	87	50.0	36	9 US-09-848-834A-16	Sequence 16, Appl
10	87	50.0	37	9 US-09-848-834A-14	Sequence 14, Appl
11	87	50.0	50	9 US-09-848-834A-18	Sequence 18, Appl
12	87	50.0	51	9 US-09-848-834A-20	Sequence 20, Appl
13	86	49.4	31	9 US-09-848-834A-9	Sequence 9, Appl
14	80	46.0	75	11 US-09-847-102A-33	Sequence 33, Appl
15	80	46.0	75	12 US-10-285-976-231	Sequence 231, Appl

16	77	44.3	16	9 US-09-848-834A-8	Sequence 8, Appl
17	72	41.4	15	11 US-09-747-802-16	Sequence 16, Appl
18	72	41.4	15	11 US-09-747-802-30	Sequence 30, Appl
19	72	41.4	15	11 US-09-865-294-8	Sequence 8, Appl
20	72	41.4	15	11 US-09-865-294-22	Sequence 22, Appl
21	72	41.4	15	12 US-10-261-446-20	Sequence 20, Appl
22	72	41.4	45	12 US-10-076-674-9	Sequence 9, Appl
23	72	41.4	550	10 US-09-873-233A-18	Sequence 18, Appl
24	72	41.4	550	10 US-09-873-233A-20	Sequence 20, Appl
25	71	40.8	17	11 US-09-305-924-4	Sequence 4, Appl
26	69	39.7	15	11 US-09-747-802-37	Sequence 37, Appl
27	69	39.7	15	11 US-09-865-294-29	Sequence 29, Appl
28	69	39.7	16	15 US-10-223-711-3	Sequence 3, Appl
29	69	39.7	19	11 US-09-747-802-48	Sequence 48, Appl
30	69	39.7	19	11 US-09-865-294-40	Sequence 40, Appl
31	66	37.9	19	11 US-09-747-802-51	Sequence 51, Appl
32	66	37.9	19	11 US-09-865-294-43	Sequence 43, Appl
33	66	37.9	662	11 US-09-951-061A-141	Sequence 141, Appl
34	64	36.8	15	11 US-09-747-802-35	Sequence 35, Appl
35	64	36.8	15	11 US-09-747-802-38	Sequence 38, Appl
36	64	36.8	15	11 US-09-747-802-40	Sequence 40, Appl
37	64	36.8	15	11 US-09-747-802-42	Sequence 42, Appl
38	64	36.8	15	11 US-09-865-294-27	Sequence 27, Appl
39	64	36.8	15	11 US-09-865-294-30	Sequence 30, Appl
40	64	36.8	15	11 US-09-865-294-32	Sequence 32, Appl
41	64	36.8	15	11 US-09-865-294-34	Sequence 34, Appl
42	64	36.8	19	11 US-09-747-802-49	Sequence 49, Appl
43	64	36.8	19	11 US-09-747-802-55	Sequence 55, Appl
44	64	36.8	19	11 US-09-865-294-38	Sequence 38, Appl
45	64	36.8	19	11 US-09-865-294-41	Sequence 41, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-848-834A-13  
Sequence 13, Application US/09848834A  
Patent No. US20020076416A1  
GENERAL INFORMATION:  
APPLICANT: Aptton Corporation  
TITLE OF INVENTION: Chimeric Peptide Immunogens  
FILE REFERENCE: 1102865-0047  
CURRENT APPLICATION NUMBER: US/09/848,834A  
CURRENT FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: 60/202,328  
PRIOR FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 13  
LENGTH: 34  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the human GnRH hormone  
OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of the Measles virus fusion protein, he Measles virus fusion protein,  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(10)  
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
NAME/KEY: PEPTIDE  
LOCATION: (11)..(18)  
OTHER INFORMATION: Spacer peptide  
NAME/KEY: PEPTIDE  
LOCATION: (19)..(34)  
OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein, F  
NAME/KEY: MOD RES  
LOCATION: (1)..(1)  
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
US-09-848-834A-13

Query Match 99.4%; Score 173; DB 9; Length 34;



C:Superfamily: parainfluenza virus cell fusion protein  
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>  
 F:89-106/Domain: transmembrane #status predicted <TM1>  
 F:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>  
 F:194-219/Domain: transmembrane #status predicted <TM2>  
 F:575-595/Domain: transmembrane #status predicted <TM3>  
 F:110,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.4%; Score 65; DB 1; Length 631;  
 Best Local Similarity 68.4%; Pred. No. 0.53;  
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLLSEIKGVVHRLEGV 33  
 Db 365 SYPTLSEVRGVVHRLEAV 383

Search completed: September 24, 2003, 17:49:32  
 Job time : 11.9612 secs

```

Db      284 LSEIKGVIVHRLSV 298
|||||
RESULT 11
VGNZCD
cell fusion glycoprotein precursor - canine distemper virus
N:Contains: fusion protein F1; fusion protein F2
C:Species: canine distemper virus
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: J50321
R:Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.
Virus Res. 8, 373-386, 1987
A:Title: The nucleotide sequence of the gene encoding the F protein of canine distemper
A:Reference number: J50321; MUID:88129050; PMID:3433924
A:Accession: J50321
A:Molecule type: mRNA
A:Residues: 1-662 <BAR>
A:Cross-references: GB:M21849; NID:g323241; PIDN:AAA42878.1; PID:g323242
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-135/Domain: signal sequence #status predicted <SIG>
F:136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>
F:225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>
F:606-629/Domain: transmembrane #status predicted <MEM>
F:62,141,173,179,517/Binding site: carbohydrate (Asn) #status predicted
Query Match 37.9%; Score 66; DB 1; Length 662;
Best Local Similarity 73.7%; Pred. No. 0.41;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy      15 SLKLLSEIKGVIVHRLGV 33
| |||:|||||
Db      396 SYPTLSEVKGVIVHRLV 414
|||||

RESULT 12
S21382
cell fusion protein - canine distemper virus
C:Species: canine distemper virus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C:Accession: S21382
R:Wild, T.F.; Bernard, A.; Spehner, D.; Villevall, D.; Drillien, R.
submitted to the EMBL Data Library, April 1992
A:Description: Vaccination of mice against canine distemper virus induced encephalitis
A:Reference number: S21382
A:Accession: S21382
A>Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-662 <WIL>
A:Cross-references: EMBL:X65509; NID:g58853; PIDN:CAA46481.1; PID:g58854
C:Superfamily: parainfluenza virus cell fusion protein
Query Match 37.9%; Score 66; DB 2; Length 662;
Best Local Similarity 73.7%; Pred. No. 0.41;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy      15 SLKLLSEIKGVIVHRLGV 33
| |||:|||||
Db      396 SYPTLSEVKGVIVHRLV 414
|||||

RESULT 13
S47034
cell fusion protein precursor - porpoise morbillivirus
N:Alternate names: F protein
C:Species: porpoise morbillivirus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C:Accession: S47034
R:Bolt, G.; Gottschalk, E.; Blixenkron-Moeller, M.; Wishaupt, R.G.A.; Welsh, M.J.; Ed
submitted to the EMBL Data Library, July 1994

```

```

A:Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbi
A:Reference number: S47034
A:Accession: S47034
A:Molecule type: mRNA
A:Residues: 1-552 <BOL>
A:Cross-references: EMBL:X80757; NID:g520639; PIDN:CAA56731.1; PID:g520640
A:Experimental source: isolate Ulster 88
A:Note: the source is designated as Cetacean morbillivirus
C:Superfamily: parainfluenza virus cell fusion protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-552/Product: fusion protein #status predicted <MAT>
Query Match 37.4%; Score 65; DB 2; Length 552;
Best Local Similarity 86.7%; Pred. No. 0.46;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy      19 LSEIKGVIVHRLGV 33
| |||:|||||
Db      290 LSEVKGVIVHRLV 304
|||||

RESULT 14
VGNZPD
cell fusion glycoprotein precursor - phocine distemper virus
N:Contains: fusion protein F1; fusion protein F2
C:Species: phocine distemper virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 25-Oct-1996
C:Accession: JQ1368
R:Koevanees, J.; Blixenkron-Moeller, M.; Sharma, B.; Oervell, C.; Norrby, E.
J. Gen. Virol. 72, 2959-2966, 1991
A:Title: The nucleotide sequence and deduced amino acid composition of the haemaggl
A:Reference number: JQ1368; MUID:92113538; PMID:1765768
A:Accession: JQ1368
A:Molecule type: genomic RNA
A:Residues: 1-631 <KOV>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F:89-106/Domain: transmembrane #status predicted <TM1>
F:189-193/Region: cleavage processing #status predicted
F:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F:194-212/Domain: transmembrane #status predicted <TM2>
F:575-595/Domain: transmembrane #status predicted <TM3>
F:110,142,148,486/Binding site: carbohydrate (Asn) #status predicted
Query Match 37.4%; Score 65; DB 1; Length 631;
Best Local Similarity 68.4%; Pred. No. 0.53;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy      15 SLKLLSEIKGVIVHRLGV 33
| |||:|||||
Db      365 SYPTLSEVKGVIVHRLV 383
|||||

RESULT 15
A48346
cell fusion glycoprotein precursor - phocine distemper virus (strain Ulster/88)
N:Contains: fusion protein F1; fusion protein F2
C:Species: phocine distemper virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 25-Oct-1996
C:Accession: A48346
R:Curran, M.D.; Lu, Y.J.; Rima, B.K.
Arch. Virol. 126, 159-169, 1992
A:Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced a
A:Reference number: A48346; MUID:92398437; PMID:1524494
A:Accession: A48346
A:Molecule type: mRNA
A:Residues: 1-631 <CUR>
A:Note: sequence extracted from NCBI backbone (NCBI:113098, NCBIP:113099)
C:Genetics:
A:Gene: F

```

C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:26-110/Product: cell fusion glycoprotein F2 #status predicted <FF2>  
 F:111-553/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
 F:501-517/Domain: transmembrane #status predicted <TMN>  
 F:32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.4%; Score 72; DB 1; Length 553;  
 Best Local Similarity 100.0%; Pred. No. 0.047;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33  
 |||||||:|||||  
 Db 291 LSEIKGVIVHRLEGV 305

RESULT 7  
 VGNZRL  
 cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)  
 N:Contains: fusion glycoprotein F1; fusion glycoprotein F2  
 C:Species: rinderpest virus  
 C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 25-Oct-1996  
 C:Accession: A31051  
 R:Hsu, D.; Yamanaka, M.; Miller, J.; Dale, B.; Grubman, M.; Yilma, T.  
 Virology 166, 149-153, 1988  
 A:Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis  
 A:Reference number: A31051; MUID:88322864; PMID:3411983  
 A:Accession: A31051  
 A:Molecule type: genomic RNA  
 A:Residues: 1-546 <HSU>  
 C:Genetics:  
 A:Gene: F  
 C:Superfamily: parainfluenza virus cell fusion protein  
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-108/Product: cell fusion glycoprotein F2 #status predicted <FF1>  
 F:109-546/Product: cell fusion glycoprotein F1 #status predicted <FF2>  
 F:109-134/Domain: transmembrane #status predicted <TM1>  
 F:491-513/Domain: transmembrane #status predicted <TM2>  
 F:25,57,63,518/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.8%; Score 71; DB 1; Length 546;  
 Best Local Similarity 93.3%; Pred. No. 0.065;  
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33  
 |||||||:|||||  
 Db 284 LSEIKGVIVHRLEGV 298

RESULT 8  
 S47305  
 gene F protein - rinderpest virus  
 C:Species: rinderpest virus  
 C:Date: 20-Oct-1994 #sequence\_revision 08-Sep-1995 #text\_change 20-Sep-1999  
 C:Accession: S47305; S47301  
 R:Baron, M.D.; Barrett, T.  
 submitted to the EMBL Data Library, March 1994  
 A:Description: The sequence of the N and L genes of Rinderpest virus, and the 50 and 30  
 A:Reference number: S47283  
 A:Accession: S47305  
 A:Molecule type: mRNA  
 A:Residues: 1-546 <BAR>  
 A:Cross-references: EMBL:Z30697; NID:g535396; PIDN:CAA03181.1; PID:g535401; EMBL:Z30700;  
 C:Superfamily: parainfluenza virus cell fusion protein  
 C:Keywords: transmembrane protein

Query Match 40.8%; Score 71; DB 2; Length 546;  
 Best Local Similarity 93.3%; Pred. No. 0.065;  
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33  
 |||||||:|||||

Db 284 LSEIKGVIVHRLEGV 298

RESULT 9  
 JQ2223  
 cell fusion protein F0 precursor - phocine distemper virus  
 N:Contains: F1 and F2 chains  
 C:Species: phocine distemper virus  
 C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 24-Nov-1999  
 C:Accession: JQ2223  
 R:Visser, I.K.G.; van der Heljden, R.W.J.; van de Bildt, M.W.G.; Kenter, M.J.H.; Oever  
 J. Gen. Virol. 74, 1989-1994, 1993  
 A:Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites  
 e virus entity.  
 A:Reference number: JQ2223; MUID:93389459; PMID:8376973  
 A:Accession: JQ2223  
 A:Molecule type: mRNA  
 A:Residues: 1-542 <VIS>  
 A:Cross-references: GB:L07075  
 A:Note: the authors translated the codon ATC for residue 4 as Leu  
 C:Comment: This fusion protein F0 is cleaved into F1 and F2 chains.  
 C:Genetics:  
 A:Gene: F  
 C:Superfamily: parainfluenza virus cell fusion protein  
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-15/Domain: signal sequence #status predicted <SIG>  
 F:16-542/Product: fusion protein #status predicted <MAT>  
 F:16-99/Product: F2 chain #status predicted <F2C>  
 F:105-542/Product: F1 chain #status predicted <FIC>  
 F:105-135/Region: hydrophobic  
 F:486-512/Domain: transmembrane #status predicted <TMW>  
 F:21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.9%; Score 66; DB 2; Length 542;  
 Best Local Similarity 73.7%; Pred. No. 0.33;  
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 LKLLSEIKGVIVHRLEGV 33  
 ||||:|||||||:|  
 Db 276 SYPLSEVKGIVVHRLEAV 294

RESULT 10  
 VGNZRL  
 cell fusion glycoprotein precursor - rinderpest virus (strain L)  
 N:Contains: fusion glycoprotein F1; fusion glycoprotein F2  
 C:Species: rinderpest virus  
 C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 16-Jul-1999  
 C:Accession: A28921  
 R:Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.  
 Virology 164, 523-530, 1988  
 A:Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of  
 A:Reference number: A28921; MUID:88219541; PMID:3285575  
 A:Accession: A28921  
 A:Molecule type: mRNA  
 A:Residues: 1-546 <TSU>  
 A:Cross-references: GB:M20870; NID:g333898; PIDN:AAA47399.1; PID:g333899  
 C:Genetics:  
 A:Gene: F  
 C:Superfamily: parainfluenza virus cell fusion protein  
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>  
 F:105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>  
 F:109-133/Domain: transmembrane #status predicted <TN1>  
 F:485-513/Domain: transmembrane #status predicted <TN2>  
 F:25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.9%; Score 66; DB 1; Length 546;  
 Best Local Similarity 93.3%; Pred. No. 0.33;  
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33

```

RESULT 3
JU0274
cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain x
N:Contains: fusion glycoprotein F1; fusion glycoprotein F2
C:Species: subacute sclerosing panencephalitis virus, SSPEV
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000
C:Accession: JU0274
R:Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.
Virus Genes 4, 173-181, 1990
A:Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe
A:Reference number: JU0274; MUID:90385702; PMID:1698327
A:Accession: JU0274
A:Molecule type: mRNA
A:Residues: 1-534 <KOW>
A:Cross-references: EMBL:D10548; NID:g222256; PIDN:BAA01405.1; PID:g222257
A:Note: the authors translated the codon GTA for residue 459 as Gly and GGG for residue
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FE2>
F:108-534/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F:498-514/Domain: transmembrane #status predicted <TMN>
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 41.4%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
Db 288 LSEIKGVIVHRLEGV 302
|||||
|||||

RESULT 4
S47300
gene F protein - rinderpest virus
C:Species: rinderpest virus
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C:Accession: S47300; PQ0865
R:Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.
submitted to the EMBL Data Library, March 1994
A:Description: The complete nucleotide sequence of the fusion protein gene of the vaccin
A:Reference number: S47299
A:Accession: S47300
A:Molecule type: DNA
A:Residues: 1-546 <EVA>
A:Cross-references: EMBL:Z31656; NID:g535406; PIDN:CAA83482.1; PID:g535407
R:Chamberlain, R.W.; Wanwayl, H.M.; Hockley, E.; Shailla, M.S.; Goatley, L.; Knowles, N.J.
J. Gen. Virol. 74, 2775-2780, 1993
A:Title: Evidence for different lineages of rinderpest virus reflecting their geographic
A:Reference number: PQ0865; MUID:94103786; PMID:827286
A:Accession: PQ0865
A:Molecule type: mRNA
A:Residues: 86-191 <CHA>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein

Query Match 41.4%; Score 72; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
Db 284 LSEIKGVIVHRLEGV 298
|||||
|||||

RESULT 5

```

```

E48556
cell fusion glycoprotein precursor - measles virus (strain AIK-C)
C:Species: measles virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
C:Accession: E48556
R:Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.
Virus Genes 7, 67-81, 1993
A:Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the AIK
A:Reference number: A48556; MUID:93227570; PMID:8470368
A:Accession: E48556
A:Molecule type: genomic RNA
A:Residues: 1-550 <MOR>
A:Cross-references: GB:S58435; NID:g299460; PIDN:AAB26145.1; PID:g299465
A:Note: sequence extracted from NCBI backbone (NCBIN:129264, NCBIP:129272)
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F:108-550/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F:113-138/Region: hydrophobic
F:495-514/Domain: transmembrane #status predicted <TMN>
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 41.4%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
Db 288 LSEIKGVIVHRLEGV 302
|||||
|||||

RESULT 6
VGN2MV
cell fusion glycoprotein precursor - measles virus
C:Species: measles virus
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1989 #text_change 16-Jun-2000
C:Accession: A26962; A25616; PQ0384
R:Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.
J. Gen. Virol. 68, 1695-1703, 1987
A:Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and co
A:Reference number: A92794; MUID:8724816; PMID:3585281
A:Accession: A26962
A:Molecule type: mRNA
A:Residues: 1-553 <BUC>
A:Cross-references: GB:D00090; NID:g222061; PIDN:BAA00056.1; PID:g222062
A:Experimental source: strain Halle
R:Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Englund, G.; Bellini
Virology 155, 508-523, 1986
A:Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles v
A:Reference number: A94350; MUID:87071668; PMID:3788062
A:Accession: A25616
A:Molecule type: mRNA
A:Residues: 4-553 <RIC>
A:Cross-references: GB:M14915; NID:g331762; PIDN:AAA46423.1; PID:g331763
A:Experimental source: strain Edmonston
R:Schulz, T.F.; Hoar, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compari
A:Reference number: PQ0374; MUID:92300360; PMID:1607874
A:Accession: PQ0380
A:Molecule type: genomic RNA
A:Residues: 272-553 <SCH1>
A:Experimental source: isolate CL
A:Accession: PQ0384
A:Molecule type: genomic RNA
A:Residues: 272-553 <SCH2>
A:Experimental source: isolate SE
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:30:06 ; Search time 11.9612 Seconds  
(without alignments)  
273.361 Million cell updates/sec

Title: US-09-848-834A-13

Perfect score: 174

Sequence: 1 XHWSYGLRPGSGPSLKLLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	41.4	282	2 P00376	cell fusion glycop
2	72	41.4	282	2 P00388	cell fusion glycop
3	72	41.4	534	1 J00274	cell fusion glycop
4	72	41.4	546	2 S47300	gene F protein - r
5	72	41.4	550	1 E48556	cell fusion glycop
6	72	41.4	553	1 VGNZMV	cell fusion glycop
7	71	40.8	546	1 VGNZRK	cell fusion glycop
8	71	40.8	546	2 J02233	gene F protein - r
9	66	37.9	542	2 J02233	cell fusion protei
10	66	37.9	546	1 VGNZRL	cell fusion glycop
11	66	37.9	662	2 VGNZCD	cell fusion glycop
12	66	37.9	662	2 S21382	cell fusion protei
13	65	37.4	552	2 S47034	cell fusion protei
14	65	37.4	631	1 VGNZPD	cell fusion glycop
15	65	37.4	631	1 A48346	cell fusion glycop
16	64	36.8	67	2 I78541	gonadoliberin prec
17	64	36.8	92	1 RHUG	gonadoliberin prec
18	62	35.6	89	2 I51423	gonadoliberin prec
19	60	34.5	546	2 S55386	cell fusion protei
20	58	33.3	10	1 RHPGG	gonadoliberin - pi
21	58	33.3	10	1 RHSHG	gonadoliberin - sh
22	58	33.3	90	1 RHMSG	gonadoliberin prec
23	58	33.3	92	1 RHRTG	gonadoliberin prec
24	56.5	32.5	98	2 I50739	gonadotropin-relea
25	56	32.2	636	2 S47299	gene F protein - r
26	55	31.6	92	2 I50644	gonadoliberin I pr
27	54.5	31.3	80	2 S39779	aldehyde reductase
28	54.5	31.3	249	2 A41497	36K antigen pra
29	54	31.0	10	1 RHAQ1	gonadoliberin I -

#### ALIGNMENTS

##### RESULT 1

P00376

cell fusion glycoprotein - measles virus (strain TT) (fragment)

C:Species: measles virus

C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 24-Nov-1999

C:Accession: P00376

R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.

J. Gen. Virol. 73, 1581-1586, 1992

A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compa

A:Reference number: P00374; MUID:92300360; PMID:1607874

A:Accession: P00376

A:Molecule type: genomic RNA

A:Residues: 1-282 <SCH>

C:Genetics:

A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein

C:Keywords: glycoprotein; membrane fusion

Query Match 41.4%; Score 72; DB 2; Length 282;

Best Local Similarity 100.0%; Pred. No. 0.022;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33

|||||

Db 20 LSEIKGVIVHRLEGV 34

##### RESULT 2

P00388

cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)

C:Species: measles virus

C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 24-Nov-1999

C:Accession: P00388

R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.

J. Gen. Virol. 73, 1581-1586, 1992

A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compa

A:Reference number: P00374; MUID:92300360; PMID:1607874

A:Accession: P00388

A:Molecule type: genomic RNA

A:Residues: 1-282 <SCH>

C:Genetics:

A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein

C:Keywords: glycoprotein; membrane fusion

Query Match 41.4%; Score 72; DB 2; Length 282;

Best Local Similarity 100.0%; Pred. No. 0.022;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33

|||||

Db 20 LSEIKGVIVHRLEGV 34

phosphatidylcholin  
hypothetical prote  
protocatechuate 3,  
D-amino acid oxida  
hypothetical prote  
probable pra prote  
hypothetical prote  
protein F5A9.22 (f  
oligopeptide ABC-t  
myo-inositol-1-pho  
hypothetical prote  
phycobilisome core  
UDP-N-acetylmurama  
hypothetical prote  
protein C13A10.3 l  
hypothetical prote



Query Match	35.68;	Score 62;	DB 1;	Length 89;
Best Local Similarity	25.4%;	Pred. No. 0.057;		
Matches 15;	Conservative	7;	Mismatches 9;	Indels 28; Gaps 1;
QY	2	HWSYGLRPGSSGSLKL	-----	LSIKGVIVHRLG 32
db	25	HWSYGLRPGGRKTESLDWMYHETPNEVALFPELRLECSVPQSLNVLRCALNMWLEG	83	

FT	CHAIN	27	115	PROTEIN F2.
FT	CHAIN	116	529	PROTEIN F1.
FT	TRANSMEM	116	139	POTENTIAL.
FT	DOMAIN	140	497	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	498	518	POTENTIAL.
FT	DOMAIN	519	529	CYTOPLASMIC (POTENTIAL).
FT	DISULFID	71	198	LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT	CARBOHYD	32	32	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	64	64	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	70	70	N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE	529 AA; 57331 MW; AE987BC9F07E9AA9	CRC64;	
Query Match 36.8%; Score 64; DB 1; Length 529;				
Best Local Similarity 93.3%; Pred. No. 0.21;				
Matches 14; Conservative 0; Mismatches 1; Indels 0;				
QY	19 LSEIKGVIVHRLEGV 33			
Db	291 LSEIKGVIVHRLEGV 305			
RESULT 14				
ID	GONL_PAGMA	STANDARD;	PRT;	95 AA.
AC	P70074;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH (LH-RH I) (Luliberin I)).			
GN	GNRH			
OS	Pagrus major (Red sea bream) (Chrysophrys major).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;			
OC	Sparidae; Pagrus.			
OX	NCBI_TaxID=143350;			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	Okuzawa K., Granneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.			
RL	Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- SIMILARITY: Belongs to the GNRH family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL out-			
CC	the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in-			
CC	modified and this statement is not removed. Usage by and for com-			
CC	entities requires a license agreement (see http://www.isb-sib.ch/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D86582; BAA13129.1; ..			
DR	InterPro; IPR002012; GNRH			
DR	InterPro; IPR004079; GonadoliberinI.			
DR	Pfam; PF00446; GNRH; 1.			
DR	PRINTS; PR01541; GONADOLIBRN1.			
DR	PROSITE; PS00473; GNRH; 1.			
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamic			
KW	Signal; Multigene family; Pyroglutamate carboxylic acid.			
FT	SIGNAL	1	23	POTENTIAL.
FT	CHAIN	24	95	PROGONADOLIBERIN I.
FT	PEPTIDE	24	33	GONADOLIBERIN I.
FT	PEPTIDE	37	95	GNRH-ASSOCIATED PEPTIDE I (POTENTIAL
FT	MOD_RES	24	24	PYROLIDONE CARBOXYLIC ACID (BY
FT	MOD_RES	33	33	SIMILARITY).
FT	MOD_RES	33	33	AMIDATION (G-34 PROVIDE AMIDE GROUP)
FT	MOD_RES	33	33	(BY SIMILARITY).
SEQ	SEQUENCE	95 AA; 10566 MW; 61E7G990328D73E	CRC64;	

```

RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female rhesus
RT macaques.";
RL Neuroendocrinology 60:346-359(1994).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S75918; AAB33096.1; -
DR PIR; I78541; I78541.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRN1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Pyrrolidone carboxylic acid.
FT NON_TER 1 1
FT SIGNAL <1 5 BY SIMILARITY.
FT CHAIN 6 >67 PROGONADOLIBERIN I.
FT PEPTIDE 6 15 GONADOLIBERIN I.
FT PEPTIDE 19 >67 GnRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY (BY SIMILARITY).
FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
FT NON_TER 67 67
FT SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 36.8%; Score 64; DB 1; Length 67;
Best Local Similarity 40.0%; Pred. No. 0 022;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSGSPSLKLLSEIKGVI 26
DB 7 HWSYGLRPGGKRDAENLMDSFQIV 31

RESULT 12
GONI_HUMAN GONI_HUMAN STANDARD; PRT; 92 AA.
ID P01148;
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
DE peptide I].
GN GnRH1 OR GnRH OR LHRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89366682; PubMed=2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing

```

```

RT hormone gene.";
RL Nucleic Acids Res. 17:6403-6403(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT SER-16.
RX MEDLINE=85012739; PubMed=6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
RT releasing hormone.";
RL Nature 311:666-668(1984).
RN [4]
RP SEQUENCE OF 24-33.
RX MEDLINE=83126573; PubMed=6760865;
RA Tan L., Rousseau P.;
RT "The chemical identity of the immunoreactive LHRH-like peptide
RT biosynthesized in the human placenta.";
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
RN [5]
RP VARIANT SER-16.
RX MEDLINE=99318093; PubMed=10391209;
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238(1999).
RN [6]
RP ERRATUM.
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RL Nat. Genet. 23:373-373(1999).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC Lutrepulse or Lutrelef (Ferring Pharmaceuticals) and Relisorm
CC (Serono).
CC -!- SIMILARITY: Belongs to the GnRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01059; CAA25526.1; -
DR EMBL; M12578; AAA35916.1; -
DR EMBL; X15215; CAA33285.1; -
DR PIR; S05308; RHHUG.
DR Genew; HGNC:4419; GnRH1.
DR MIM; 152760; -
DR GO; GO:0005625; C:soluble fraction; TAS.
DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.
DR GO; GO:0007267; P:cell-cell signaling; TAS.
DR GO; GO:0007275; P:development; TAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.

```

CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M21514; AAA47400.1; -  
 CC PIR; A31051; VGNZK.  
 CC HSSP; P04849; LSVF.  
 CC InterPro; IPR000776; Fusion\_gly.  
 CC Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
 CC KW SIGNAL 1 19  
 CC CHAIN 20 546 FUSION GLYCOPROTEIN F0.  
 CC CHAIN 20 108 F2 PROTEIN.  
 CC CHAIN 109 546 F1 PROTEIN.  
 CC DOMAIN 104 108 ARG-RICH (BASIC).  
 CC TRANSMEM 109 133 POTENTIAL.  
 CC TRANSMEM 484 513 POTENTIAL.  
 CC DOMAIN 514 517 ARG/LYS-RICH (BASIC).  
 CC DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
 CC CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SEQUENCE 546 AA; 58662 MW; 476D74DC18BCFCF CRC64;  
 CC  
 CC Query Match 37.48; Score 65; DB 1; Length 546;  
 CC Best Local Similarity 86.78; Pred. No. 0.15;  
 CC Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC Qy 19 LSEIKGVIVHRLGV 33  
 CC |||||:|||||  
 CC Db 284 LSEIKGVIIHRLGV 298  
 CC  
 CC RESULT 10  
 CC VGLF\_PHODV STANDARD; PRT; 631 AA.  
 CC AC P2886;  
 CC DT 01-DEC-1992 (Rel. 24, Created)  
 CC DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
 CC DE Fusion glycoprotein F1].  
 CC GN F.  
 CC OS Phocine distemper virus (PDV).  
 CC OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 CC OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 CC OX NCBI\_TaxID=11240;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=U1ster/88;  
 CC RX MEDLINE=92113536; PubMed=1765768;  
 CC RA Koovamees J., Blixenkrone-Moeller M., Sharma B., Oervell C.,  
 CC RA Norby E.;  
 CC RT "The nucleotide sequence and deduced amino acid composition of the  
 CC RT haemagglutinin and fusion proteins of the morbillivirus phocid  
 CC RT distemper virus.";  
 CC RL J. Gen. Virol. 72:2959-2966(1991).  
 CC RN [2]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=U1ster/88;  
 CC RX MEDLINE=92398437; PubMed=1524494;  
 CC RA Curran M.D., Lu Y.J., Rima B.K.;  
 CC RT "The fusion protein gene of phocine distemper virus: nucleotide and  
 CC RT deduced amino acid sequences and a comparison of morbillivirus fusion  
 CC RT proteins";  
 CC RL Arch. Virol. 126:159-169(1992).

RN [3]  
 RP SEQUENCE OF 95-631 FROM N.A.  
 RC STRAIN=U1ster/88;  
 RX MEDLINE=91089508; PubMed=2264246;  
 RA Curran M.D., Loan D.O., Rima B.K., Kennedy S.;  
 RT "Nucleotide sequence analysis of phocine distemper virus reveals its  
 RT distinctness from canine distemper virus.";  
 RL Vet. Rec. 127:430-431(1990).  
 CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
 CC MEMBRANES.  
 CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
 CC LINKED BY A DISULFIDE BOND. PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
 CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D10371; BAA01206.1; -  
 CC PIR; A48346; A48346.  
 CC PIR; JQ1368; VGNZPD.  
 CC HSSP; P04849; LSVF.  
 CC InterPro; IPR000776; Fusion\_gly.  
 CC Pfam; PF00523; Fusion\_gly; 1.  
 CC Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
 CC KW SIGNAL 1 1  
 CC CHAIN 1 631 FUSION GLYCOPROTEIN F0.  
 CC CHAIN 194 631 F2 PROTEIN.  
 CC DISULFID 149 276 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
 CC TRANSMEM 89 106 POTENTIAL.  
 CC TRANSMEM 194 212 POTENTIAL.  
 CC TRANSMEM 575 595 POTENTIAL.  
 CC CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 63 63 I -> V (IN REF. 2).  
 CC SEQUENCE 631 AA; 68873 MW; D1FC87CDD426E9B8 CRC64;  
 CC  
 CC Query Match 37.48; Score 65; DB 1; Length 631;  
 CC Best Local Similarity 68.48; Pred. No. 0.18;  
 CC Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 CC  
 CC Qy 15 SLKLLSEIKGVIVHRLGV 33  
 CC | :|||:|||||  
 CC Db 365 SYPTLSEVKGVVHRLGV 383  
 CC  
 CC RESULT 11  
 CC GONI\_MACMU STANDARD; PRT; 67 AA.  
 CC ID GONI\_MACMU  
 CC AC P55247;  
 CC DT 01-OCT-1996 (Rel. 34, Created)  
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LH-RH I)  
 CC DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 CC DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]  
 CC DE (Fragment).  
 CC GN GNRH1 OR GNRH OR LHRH.  
 CC OS Macaca mulatta (Rhesus macaque).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 CC OC Cercopithecoidea; Macaca.  
 CC OX NCBI\_TaxID=9544;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC TISSUE=Hypothalamus;



SEQUENCE FROM N.A.  
MEDLINE=95088609; PubMed=7996154;  
Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;  
"Nucleotide sequence comparisons of the fusion protein gene from  
rinderpest virus and attenuated strains of rinderpest virus.";  
J. Gen. Virol. 75:3611-3617(1994).  
-!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
MEMBRANES.  
-!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
LINKED BY A DISULFIDE BOND.  
-!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; Z30700; CAA83186.1; -;  
EMBL; Z30697; CAA83181.1; -;  
PIR; S47305; S47305.  
HSSP; P04849; 1SVF.  
InterPro; IPR000776; Fusion\_gly.  
Pfam; PF00523; fusion\_gly; 1.  
Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
SIGNAL 1 19  
CHAIN 20 546 FUSION GLYCOPROTEIN F0.  
CHAIN 20 108 F2 PROTEIN.  
CHAIN 109 546 F1 PROTEIN.  
DOMAIN 104 108 ARG/LYS-RICH (BASIC).  
TRANSMEM 109 133 POTENTIAL.  
TRANSMEM 484 513 POTENTIAL.  
DOMAIN 514 517 ARG/LYS-RICH (BASIC).  
DOMAIN 514 517 ARG/LYS-RICH (BASIC).  
DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SEQUENCE 546 AA; 58705 MW; ED3DF8AFDEBCEB95 CRC64;  
Query Match 40.8%; Score 71; DB 1; Length 546;  
Best Local Similarity 93.3%; Pred. No. 0.021;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 19 LSEIKGVIVHRLGV 33  
DB 284 LSEIKGVIVHRLGV 298  
RESULT 6  
VGLF\_RINDL STANDARD; PRT; 546 AA.  
AC P10864;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-JUL-1989 (Rel. 11, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
Fusion glycoprotein F1].  
GN F.  
OS Rinderpest virus (strain L) (RDV).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11243;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88219541; PubMed=3285575;  
RA Tsukiyama K., Yoshikawa Y., Yamanouchi K.;  
"Fusion glycoprotein (F) of rinderpest virus: entire nucleotide  
sequence of the F mRNA, and several features of the F protein.";  
Virology 164:523-530(1988).  
-----  
-!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
MEMBRANES.  
-!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
LINKED BY A DISULFIDE BOND.  
-!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M20870; AAA47399.1; -;  
PIR; A28921; VGNZRL.  
HSSP; P04849; 1SVF.  
InterPro; IPR000776; Fusion\_gly.  
Pfam; PF00523; fusion\_gly; 1.  
Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
SIGNAL 1 19  
CHAIN 20 546 FUSION GLYCOPROTEIN F0.  
CHAIN 20 108 F2 PROTEIN.  
CHAIN 109 546 F1 PROTEIN.  
DOMAIN 104 108 ARG/LYS-RICH (BASIC).  
TRANSMEM 109 133 POTENTIAL.  
TRANSMEM 484 513 POTENTIAL.  
DOMAIN 514 517 ARG/LYS-RICH (BASIC).  
DOMAIN 514 517 ARG/LYS-RICH (BASIC).  
DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SEQUENCE 546 AA; 58911 MW; 985029418F28FFB5 CRC64;  
Query Match 37.9%; Score 66; DB 1; Length 546;  
Best Local Similarity 93.3%; Pred. No. 0.11;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 19 LSEIKGVIVHRLGV 33  
DB 284 LSEIKGVIVHRLGV 298  
RESULT 7  
VGLF\_CDVO STANDARD; PRT; 662 AA.  
AC P12569; Q65991;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
Fusion glycoprotein F1].  
GN F.  
OS Canine distemper virus (strain Onderstepoort) (CDV).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11233;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88129050; PubMed=3433924;  
RA Barrett T., Clarke D.K., Evans S.A., Rima B.K.;  
"The nucleotide sequence of the gene encoding the F protein of canine  
distemper virus: a comparison of the deduced amino acid sequence with  
other paramyxoviruses.";  
Virus Res. 8:373-386(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93227696; PubMed=8470428;  
RA Wild T.F., Bernard A., Spehner D., Villevall D., Drillien R.;  
"Vaccination of mice against canine distemper virus-induced  
encephalitis with vaccinia virus recombinants encoding measles or  
canine distemper virus antigens.";

SEQUENCE FROM N.A.  
MEDLINE=95088609; PubMed=7996154;  
Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;  
"Nucleotide sequence comparisons of the fusion protein gene from  
rinderpest virus and attenuated strains of rinderpest virus.";  
J. Gen. Virol. 75:3611-3617(1994).  
-!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR  
MEMBRANES.  
-!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2  
LINKED BY A DISULFIDE BOND.  
-!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN  
FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M20870; AAA47399.1; -;  
PIR; A28921; VGNZRL.  
HSSP; P04849; 1SVF.  
InterPro; IPR000776; Fusion\_gly.  
Pfam; PF00523; fusion\_gly; 1.  
Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.  
SIGNAL 1 19  
CHAIN 20 546 FUSION GLYCOPROTEIN F0.  
CHAIN 20 108 F2 PROTEIN.  
CHAIN 109 546 F1 PROTEIN.  
DOMAIN 104 108 ARG/LYS-RICH (BASIC).  
TRANSMEM 109 133 POTENTIAL.  
TRANSMEM 484 513 POTENTIAL.  
DOMAIN 514 517 ARG/LYS-RICH (BASIC).  
DOMAIN 514 517 ARG/LYS-RICH (BASIC).  
DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).  
CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SEQUENCE 546 AA; 58911 MW; 985029418F28FFB5 CRC64;  
Query Match 37.9%; Score 66; DB 1; Length 546;  
Best Local Similarity 93.3%; Pred. No. 0.11;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 19 LSEIKGVIVHRLGV 33  
DB 284 LSEIKGVIVHRLGV 298  
RESULT 7  
VGLF\_CDVO STANDARD; PRT; 662 AA.  
AC P12569; Q65991;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
Fusion glycoprotein F1].  
GN F.  
OS Canine distemper virus (strain Onderstepoort) (CDV).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11233;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88129050; PubMed=3433924;  
RA Barrett T., Clarke D.K., Evans S.A., Rima B.K.;  
"The nucleotide sequence of the gene encoding the F protein of canine  
distemper virus: a comparison of the deduced amino acid sequence with  
other paramyxoviruses.";  
Virus Res. 8:373-386(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93227696; PubMed=8470428;  
RA Wild T.F., Bernard A., Spehner D., Villevall D., Drillien R.;  
"Vaccination of mice against canine distemper virus-induced  
encephalitis with vaccinia virus recombinants encoding measles or  
canine distemper virus antigens.";

```

DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Measles virus (strain Edmonston) (Subacute sclerose panencephalitis
OS virus),
OS Measles virus (strain Halle) (Subacute sclerose panencephalitis
OS virus),
OS Measles virus (strain Leningrad-16) (Subacute sclerose panencephalitis
OS virus),
OS Measles virus (strain Edmonston-Zagreb) (Subacute sclerose
OS panencephalitis virus),
OS Measles virus (strain Philadelphia-26) (Subacute sclerose
OS panencephalitis virus), and
OS Measles virus (strain Edmonston B) (Subacute sclerose panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11235, 11236, 70147, 70149, 70148, 70146;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Edmonston;
RX MEDLINE=87071668; PubMed=3788062;
RA Richardson C.D., Hull D., Greer P., Hasel K., Berkovich A.,
RA Englund G., Bellini W.J., Rima B., Lazzarini R.A.;
RT "The nucleotide sequence of the mRNA encoding the fusion protein of
RT measles virus (Edmonston strain): a comparison of fusion proteins
RT from several different paramyxoviruses.";
RT Virology 155:508-523(1986).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Halle;
RC MEDLINE=87224816; PubMed=3585281;
RX Buckland R., Gerald C., Barker R., Wild T.F.;
RA "Fusion glycoprotein of measles virus: nucleotide sequence of the
RA gene and comparison with other paramyxoviruses.";
RL J. Gen. Virol. 68:1695-1703(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Edmonston;
RX MEDLINE=90085790; PubMed=2596022;
RA Cattaneo R., Schmid A., Spielhofer P., Kaelin K., Bacsko K.,
RA Meulen V., Pardowitz J., Flanagan S., Rima B.K., Udem S.A.;
RT "Mutated and hypermutated genes of persistent measles viruses which
RT caused lethal human brain diseases.";
RL Virology 173:415-425(1989).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Edmonston;
RX MEDLINE=92263801; PubMed=1585658;
RA Schmid A., Spielhofer P., Cattaneo R., Bacsko K., Ter Meulen V.,
RA Billeter M.A.;
RT "Subacute sclerosing panencephalitis is typically characterized by
RT alterations in the fusion protein cytoplasmic domain of the
RT persisting measles virus.";
RL Virology 188:910-915(1992).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Edmonston, Leningrad-16, and Edmonston-Zagreb;
RX MEDLINE=94249283; PubMed=8191786;
RA Rota J.S., Wang Z.D., Rota P.A., Bellini W.J.;
RT "Comparison of sequences of the H, F, and N coding genes of measles
RT virus vaccine strains.";
RL Virus Res. 31:317-330(1994).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Philadelphia-26;
RX MEDLINE=94303181; PubMed=8030232;
RA Hummel K.B., Vanchiere J.A., Bellini W.J.;
RT "Restriction of fusion protein mRNA as a mechanism of measles virus
RT persistence.";

```

---

```

RL Virology 202:665-672(1994).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=Edmonston B;
RA Billeter M.A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14915; AAA6423.1; -
DR EMBL; X05597; CAA29090.1; ALT_INIT.
DR EMBL; K01711; AAA75498.1; ALT_INIT.
DR EMBL; K01711; AAA75499.1; -
DR EMBL; U03657; AAA56647.1; ALT_INIT.
DR EMBL; U03659; AAA56649.1; ALT_INIT.
DR EMBL; U03670; AAA56660.1; ALT_INIT.
DR EMBL; U08416; AAA50550.1; ALT_INIT.
DR EMBL; Z66517; CAA91367.1; ALT_INIT.
DR EMBL; Z66517; CAA91368.1; -
DR HSSP; P04849; LSVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 550 FUSION GLYCOPROTEIN FO.
FT CHAIN 24 112 PROTEIN F2.
FT CHAIN 113 550 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
FT DISULFD 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 550 AA; 59532 MW; 7AA4F1CA82169093 CRC64;
Query Match 41.4%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 LSEIKGVIVHRLEGV 33
| | | | | | | | | | | | | | | | | | | | |
DB 288 LSEIKGVIVHRLEGV 302
RESULT 5
VGLF_RINDR
ID VGLF_RINDR STANDARD; PRT; 546 AA.
AC P41356;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain R80K) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=36409;
RN [1]

```

```
Query Match 41.4%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVIVHRLGV 33
    |||||
Db 288 LSEIKGVIVHRLGV 302
    |||||

RESULT 2
VGLF_RINDB STANDARD; PRT; 546 AA.
AC P41360;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain RB1) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=39007;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088609; PubMed=7996154;
RA Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT "Nucleotide sequence comparisons of the fusion protein gene from
RT virulent and attenuated strains of rinderpest virus.";
RL J. Gen. Virol. 75:3611-3617(1994).
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z31656; CAA83482.1; .
DR PIR; S47300; S47300.
DR HSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
DR Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 O-LINKED (POTENTIAL).
SQ SEQUENCE 546 AA; 58418 MW; 38B539B9344F401 CRC64;

Query Match 41.4%; Score 72; DB 1; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVIVHRLGV 33
    |||||
Db 288 LSEIKGVIVHRLGV 302
    |||||

RESULT 4
VGLF_MEASE STANDARD; PRT; 550 AA.
AC P08300;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Measles virus (strain AIK-C) (Subacute sclerosing panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=36408;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227570; PubMed=8470368;
RA Mori T., Sasaki K., Hashimoto H., Makino S.;
RT "Molecular cloning and complete nucleotide sequence of genomic RNA of
RT the AIK-C strain of attenuated measles virus.";
RL Virus Genes 7:67-81(1993).
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S58435; AAB26145.1; .
DR PIR; E48556; E48556.
DR HSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
DR Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.
FT CHAIN 24 112 PROTEIN F2.
FT CHAIN 113 550 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 550 POTENTIAL.
FT DISULFID 68 195 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 29 29 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 550 AA; 59540 MW; AAC4DAB92DE0D938 CRC64;

Query Match 41.4%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVIVHRLGV 33
    |||||
Db 288 LSEIKGVIVHRLGV 302
    |||||

RESULT 4
VGLF_MEASE STANDARD; PRT; 550 AA.
AC P08300;
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:00:31 ; Search time 6.31025 Seconds  
(without alignments)  
253.383 Million cell updates/sec

Title: US-09-848-834A-13  
Perfect score: 174  
Sequence: 1 XHWSYGLRPGSGPSLKLLSEIKGVIVHRLGVE 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	72	41.4	534	1 VGLF_MEASY
2	72	41.4	546	1 VGLF_RINDB
3	72	41.4	550	1 VGLF_MEASA
4	72	41.4	550	1 VGLF_MEASE
5	71	40.8	546	1 VGLF_RINDR
6	66	37.9	546	1 VGLF_RINDL
7	66	37.9	662	1 VGLF_CDVO
8	65.5	37.6	95	1 GONL_SPAAU
9	65	37.4	546	1 VGLF_RINDK
10	65	37.4	631	1 VGLF_PHODV
11	64	36.8	67	1 GONL_MACMU
12	64	36.8	92	1 GONL_HUMAN
13	64	36.8	529	1 VGLF_MEASI
14	62.5	35.9	95	1 GONL_PAGNA
15	62	35.6	89	1 GONL_XENLA
16	61.5	35.3	61	1 GONL_SHEEP
17	60	34.5	92	1 GONL_TUPGB
18	58	33.3	63	1 GONL_MESAU
19	58	33.3	90	1 GONL_MOUSE
20	58	33.3	91	1 GONL_PIG
21	58	33.3	92	1 GONL_RAT
22	58	33.3	99	1 GONL_DICLA
23	57	32.8	95	1 GONL_MORSA
24	56.5	32.5	94	1 GONL_HAPBU
25	55	31.6	92	1 GONL_CHICK
26	54.5	31.3	249	1 PRA_MYCLE
27	54	31.0	10	1 GONL_ALLMI
28	52	29.9	213	1 PCT_BOVIN
29	51.5	29.6	74	1 GON3_ONCMY
30	51.5	29.6	90	1 GON8_RANDY
31	51	29.3	393	1 FEZ1_RAT
32	50.5	29.0	240	1 PRA_MYCTU
33	50	28.7	1131	1 APCE_ANASP

34	50	28.7	2114	1 MY9B_MOUSE	Q9qy06 mus musculus
35	49.5	28.4	342	1 MURB_ECOLI	P08373 escherichia
36	49	28.2	92	1 GON1_CAVPO	O54713 cavia porce
37	49	28.2	521	1 UBF3_HUMAN	Q9y614 homo sapien
38	48.5	27.9	359	1 PTN7_RAT	P49445 ratus norv
39	48.5	27.9	520	1 UBF3_MOUSE	Q91w36 mus musculus
40	48	27.6	80	1 GON1_CLAGA	P33439 claria gar
41	48	27.6	336	1 UL134_EBV	P03185 epstein-bar
42	48	27.6	360	1 PTN7_HUMAN	P35236 homo sapien
43	48	27.6	395	1 PTP_ENCCU	O76942 oncephalito
44	47.5	27.3	74	1 GON3_ONCTS	Q92097 oncorhynch
45	47.5	27.3	82	1 GON3_SALSA	P35629 salmo salar

ALIGNMENTS

RESULT 1

VGLF\_MEASY

ID VGLF\_MEASY STANDARD; PRT; 534 AA.

AC P26032;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;

DE Fusion glycoprotein F1].

DE

GN F.

OS Measles virus (strain Yamagata-1) (Subacute sclerosing panencephalitis virus).

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.

OX NCBI\_TaxID=11239;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90385702; PubMed=1698327;

RA Komase K., Haga T., Yoshikawa Y., Sato T.A., Yamanouchi K.;

RT "Molecular analysis of structural protein genes of the Yamagata-1 strain of defective subacute sclerosing panencephalitis virus. IV. Nucleotide sequence of the fusion gene.";

RT

RL Virus Genes 4:173-181(1990).

CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR MEMBRANES.

CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2 LINKED BY A DISULFIDE BOND.

CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL: D10548; BAA01405.1; -.

CC HSP: P04849; 1SVF.

CC InterPro: IPR000776; Fusion\_gly.

CC Pfam: PF00523; fusion\_gly; 1.

CC Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.

CC SIGNAL 1 23

CC CHAIN 24 534 FUSION GLYCOPROTEIN F0.

CC CHAIN 24 112 PROTEIN F2.

CC CHAIN 113 534 PROTEIN F1.

CC TRANSMEM 113 136 POTENTIAL.

CC DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).

CC TRANSMEM 495 515 POTENTIAL.

CC DOMAIN 516 534 CYTOPLASMIC (POTENTIAL).

CC DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).

CC CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SEQUENCE 534 AA: 57963 MW: F5B21757E643844D CRC64;



DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)  
 DE Fusion protein.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9301V;  
 RX MEDLINE=98440529; PubMed=9765410;  
 RA Takeda M., Kato A., Kobune F., Sakata H., Li Y., Shioda T., Sakai Y.,  
 RA Asakawa M., Nagai Y.;  
 RT "Measles virus attenuation associated with transcriptional impediment  
 RT and a few amino acid changes in the polymerase and accessory  
 RT proteins";  
 RL J. Virol. 72:8690-8696(1998).  
 DR EMBL; AB012949; BAA33877.1; -;  
 DR EMBL; AB012948; BAA33871.1; -;  
 DR HSP; P04849; 1SVF.  
 DR InterPro: IPR000776; Fusion\_gly.  
 DR Pfam: PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59512 MW; 7AA4F1D117197BF9 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33  
 |||||  
 Db 288 LSEIKGVIVHRLEGV 302

## RESULT 14

QSQEX1  
 ID QSQEX1 PRELIMINARY; PRT; 550 AA.  
 AC QSQEX1;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE Fusion protein.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Masusako;  
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,  
 RA Ogura H.;  
 RT "Nucleotide sequences of the fusion protein gene of subacute  
 RT sclerosing panencephalitis viruses: deduced amino acid sequences  
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or  
 RT predicted secondary structure changed.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF179430; AAF02695.1; -;  
 DR HSP; P04849; 1SVF.  
 DR InterPro: IPR000776; Fusion\_gly.  
 DR Pfam: PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59559 MW; 609EE024A7E59C54 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33  
 |||||  
 Db 288 LSEIKGVIVHRLEGV 302

## RESULT 15

QSQEW8  
 ID QSQEW8 PRELIMINARY; PRT; 550 AA.

AC QSQEW8;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OSA-2;  
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,  
 RA Ogura H.;  
 RT "Nucleotide sequences of the fusion protein gene of subacute  
 RT sclerosing panencephalitis viruses: deduced amino acid sequences  
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or  
 RT predicted secondary structure changed.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF179437; AAF02702.1; -;  
 DR HSP; P04849; 1SVF.  
 DR InterPro: IPR000776; Fusion\_gly.  
 DR Pfam: PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59315 MW; 086E51FED235EBBA CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33  
 |||||  
 Db 288 LSEIKGVIVHRLEGV 302

Search completed: September 24, 2003, 17:47:19  
 Job time : 33.964 secs



RT fusogenicity of measles virus.";  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.  
 [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NAGAHATA;  
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,  
 RA Ogura H.;  
 RT "Nucleotide sequences of the fusion protein gene of subacute  
 RT sclerosing panencephalitis viruses: deduced amino acid sequences  
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or  
 RT predicted secondary structure changed.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; D63926; BAA09958.1; -;  
 DR EMBL; AF179431; AAF02696.1; -;  
 DR HSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59530 MW; 97C991C7E2169839 CRC64;  
 Query Match 41.4%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 19 LSEIKGVIVHRLEGV 33  
 Db 288 LSEIKGVIVHRLEGV 302  
 RESULT 6  
 Q9QEX0 PRELIMINARY; PRT; 550 AA.  
 AC Q9QEX0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RC STRAIN=Toyoshima;  
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,  
 RA Ogura H.;  
 RT "Nucleotide sequences of the fusion protein gene of subacute  
 RT sclerosing panencephalitis viruses: deduced amino acid sequences  
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or  
 RT predicted secondary structure changed.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF179432; AAF02697.1; -;  
 DR HSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59504 MW; 2AA969D37FA5CA17 CRC64;  
 Query Match 41.4%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 19 LSEIKGVIVHRLEGV 33  
 Db 288 LSEIKGVIVHRLEGV 302  
 RESULT 7  
 Q9QEW9 PRELIMINARY; PRT; 550 AA.  
 AC Q9QEW9;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Fusion protein.

OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RC STRAIN=FROM N.A.  
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,  
 RA Ogura H.;  
 RT "Nucleotide sequences of the fusion protein gene of subacute  
 RT sclerosing panencephalitis viruses: deduced amino acid sequences  
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or  
 RT predicted secondary structure changed.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF179436; AAF02701.1; -;  
 DR HSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59405 MW; 0AF6DBFC5DD22BBA CRC64;  
 Query Match 41.4%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 19 LSEIKGVIVHRLEGV 33  
 Db 288 LSEIKGVIVHRLEGV 302  
 RESULT 8  
 P90330 PRELIMINARY; PRT; 550 AA.  
 ID P90330;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RC STRAIN=Nagahata(HB);  
 RA Sheng J., Watanabe M., Ueda S.;  
 RT "Selection of a neurotropic variant of measles virus.";  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RC STRAIN=FROM N.A.  
 RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;  
 RT "An amino acid alteration of F protein responsible for the enhanced  
 RT fusogenicity of measles virus.";  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; D63924; BAA09951.1; -;  
 DR HSP; P04849; 1SVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 550 AA; 59589 MW; 73E7BD457ABA39B7 CRC64;  
 Query Match 41.4%; Score 72; DB 12; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 19 LSEIKGVIVHRLEGV 33  
 Db 288 LSEIKGVIVHRLEGV 302  
 RESULT 9  
 Q9QEW7 PRELIMINARY; PRT; 550 AA.  
 ID Q9QEW7

## RESULT 2

Q04242 ID Q04242 PRELIMINARY; PRT; 537 AA.  
 AC Q04242;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89003063; PubMed=3167982;  
 RA Cattaneo R., Schmid A., Eschle D., Bacsko K., ter Meulen V.,  
 RA Billeter M.A.;  
 RT "Biased hypermutation and other genetic changes in defective measles  
 RT viruses in human brain infections.";  
 RL Cell 55:255-265(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Cattaneo R., Billeter M.A.;  
 RL Virology 0:0-0(0).  
 DR EMBL; X16567; CAA34574.1; -;  
 DR EMBL; X16567; CAA34575.1; -;  
 DR HSP; P04849; ISVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 537 AA; 58275 MW; D0A60AC66D979E06 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 537;  
 Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33  
 |||||  
 Db 291 LSEIKGVIVHRLEGV 305

## RESULT 3

Q9PXA4 ID Q9PXA4 PRELIMINARY; PRT; 545 AA.  
 AC Q9PXA4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OSA-3;  
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,  
 RA Ogura H.;  
 RT "Nucleotide sequences of the fusion protein gene of subacute  
 RT sclerosing panencephalitis viruses: deduced amino acid sequences  
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or  
 RT predicted secondary structure changed.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF179440; AAF02705.1; -;  
 DR EMBL; AF179439; AAF02704.1; -;  
 DR HSP; P04849; ISVF.  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 545 AA; 58907 MW; 0234C28AE193E77D CRC64;

Query Match 41.4%; Score 72; DB 12; Length 545;

Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33  
 |||||  
 Db 288 LSEIKGVIVHRLEGV 302

## RESULT 4

Q91HA5 ID Q91HA5 PRELIMINARY; PRT; 546 AA.  
 AC Q91HA5;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Rinderpest virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11241;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K;  
 RX MEDLINE=21014265; PubMed=11186456;  
 RA Alanot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,  
 RA Gusev A.A.;  
 RT "Primary structure of the F-gene from Rinderpest virus strain K.";  
 RL Mol. Gen. Microbiol. Virusol. 4:29-33(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K;  
 RA Alanot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,  
 RA Gusev A.A.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY035887; AAK63190.1; -;  
 DR InterPro; IPR000776; Fusion\_gly.  
 DR Pfam; PF00523; fusion\_gly; 1.  
 SQ SEQUENCE 546 AA; 58572 MW; 449B2B2DD7405F0B CRC64;

Query Match 41.4%; Score 72; DB 12; Length 546;  
 Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33  
 |||||  
 Db 284 LSEIKGVIVHRLEGV 298

## RESULT 5

P90331 ID P90331 PRELIMINARY; PRT; 550 AA.  
 AC P90331;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Fusion protein.  
 GN F.  
 OS Measles virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NAGAHATA;  
 RA Sheng J., Watanabe M., Ueda S.;  
 RT "Selection of a neurotropic variant of measles virus.";  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NAGAHATA;  
 RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;  
 RT "An amino acid alteration of F protein responsible for the enhanced

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:29:11 ; Search time 32.964 Seconds  
(without alignments)  
266.163 Million cell updates/sec

Title: US-09-848-834A-13  
Perfect score: 174  
Sequence: 1 XHWSYGLRPGSSPSLKLSEIKGIVHRLGV 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phase:\*
- 10: sp.plant:\*
- 11: sp.rodent:\*
- 12: sp.virus:\*
- 13: sp.vertibrate:\*
- 14: sp.unclassified:\*
- 15: sp.rvirus:\*
- 16: sp.bacteriap:\*
- 17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	72	41.4	534	12 Q04243	Q04243 measles vir
2	72	41.4	537	12 Q04242	Q04242 measles vir
3	72	41.4	545	12 Q9PXA4	Q9PXA4 measles vir
4	72	41.4	546	12 Q9LHA5	Q9LHA5 rinderpest
5	72	41.4	550	12 P90331	P90331 measles vir
6	72	41.4	550	12 Q9QEX0	Q9QEX0 measles vir
7	72	41.4	550	12 Q9QEW9	Q9QEW9 measles vir
8	72	41.4	550	12 P90330	P90330 measles vir
9	72	41.4	550	12 Q9QEW7	Q9QEW7 measles vir
10	72	41.4	550	12 Q9WMK4	Q9WMK4 measles vir
11	72	41.4	550	12 Q89495	Q89495 measles vir
12	72	41.4	550	12 Q8V049	Q8V049 measles vir
13	72	41.4	550	12 Q9YJ94	Q9YJ94 measles vir
14	72	41.4	550	12 Q9QEX1	Q9QEX1 measles vir
15	72	41.4	550	12 Q9QEW8	Q9QEW8 measles vir
16	72	41.4	553	12 Q93055	Q93055 measles vir

17	72	41.4	553	12 Q9IC36	Q9IC36 measles vir
18	72	41.4	553	12 P88973	P88973 measles vir
19	72	41.4	553	12 Q83536	Q83536 measles vir
20	72	41.4	553	12 O11383	O11383 measles vir
21	72	41.4	553	12 Q9IFK2	Q9IFK2 measles vir
22	72	41.4	553	12 Q83533	Q83533 measles vir
23	72	41.4	553	12 Q83525	Q83525 measles vir
24	72	41.4	553	12 Q83518	Q83518 measles vir
25	72	41.4	553	12 P88974	P88974 measles vir
26	72	41.4	553	12 Q83527	Q83527 measles vir
27	72	41.4	553	12 Q83521	Q83521 measles vir
28	72	41.4	553	12 Q83530	Q83530 measles vir
29	72	41.4	553	12 Q91248	Q91248 measles vir
30	72	41.4	553	12 Q91QP2	Q91QP2 measles vir
31	72	41.4	553	12 Q04244	Q04244 measles vir
32	72	41.4	579	12 Q9PWU4	Q9PWU4 measles vir
33	68	39.1	545	12 Q9QEW6	Q9QEW6 measles vir
34	68	39.1	553	12 O11380	O11380 measles vir
35	66	37.9	528	12 Q9YJW9	Q9YJW9 canine dist
36	66	37.9	530	12 Q8OV06	Q8OV06 canine dist
37	66	37.9	662	12 Q9DXZ2	Q9DXZ2 canine dist
38	66	37.9	662	12 Q91KN3	Q91KN3 canine dist
39	66	37.9	662	12 Q9YKL7	Q9YKL7 canine dist
40	66	37.9	662	12 Q89327	Q89327 canine dist
41	65.5	37.6	87	13 Q9YI26	Q9YI26 sparus aura
42	65	37.4	552	12 Q66147	Q66147 cetacean mo
43	64.5	37.1	68	13 Q8JIF4	Q8JIF4 acanthopagr
44	62.5	35.9	64	13 Q8JIF2	Q8JIF2 pagrus majo
45	62	35.6	552	12 Q66409	Q66409 dolphin mor

ALIGNMENTS

RESULT 1

Q04243	ID	Q04243	PRELIMINARY;	PRT;	534 AA.
AC	Q04243;				
DT	01-NOV-1996	(TReMBLrel. 01, Created)			
DT	01-NOV-1996	(TReMBLrel. 01, Last sequence update)			
DT	01-JUN-2001	(TReMBLrel. 17, Last annotation update)			
DE	Fusion protein.				
GN	F.				
OS	Measles virus.				
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;				
OC	Paramyxoviridae; Paramyxovirinae; Morbillivirus.				
OX	NCBI_TaxID=11234;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89003063; PubMed=3167982;				
RA	Cattaneo R., Schmid A., Eschle D., Bacsko K., ter Meulen V.,				
RA	Billeter M.A.;				
RT	"Biased hypermutation and other genetic changes in defective measles				
RT	viruses in human brain infections.";				
RL	Cell 55:255-265(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Cattaneo R., Billeter M.A.;				
RL	Virology 0:0-0(0).				
DR	EMBL; X16568; CAA34581.1; -				
DR	EMBL; X16568; CAA34582.1; -				
DR	HSSP; P04849; ISVF.				
DR	InterPro; IPR000776; Fusion_gly.				
DR	Pfam; PF005523; fusion_gly; 1.				
SQ	SEQUENCE 534 AA; 57899 MW; 637245E23B5BE044 CRC64;				
Query Match 41.4%; Score 72; DB 12; Length 534;					
Best Local Similarity 100.0%; Pred. No. 0.11;					
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	19	LSEIKGIVHRLGV	33		
Db	291	LSEIKGIVHRLGV	305		

Db :|||||  
506 MFNNTVSEWLRVPKVSASHLE 527

Search completed: September 24, 2003, 17:00:24  
Job time : 12.0668 secs

QY 16 LFNNFTVSFWLRVPKVSASHLE 37  
:|||||  
Db 491 MFNNFTVSFWLRVPKVSASHLE 512

RESULT 11  
US-08-913-880C-14  
; Sequence 14, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913,880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 14  
; LENGTH: 862  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 454 to 1315 of SEQ ID NO: 1  
US-08-913-880C-14

Query Match 57.0%; Score 114; DB 4; Length 862;  
Best Local Similarity 95.5%; Pred. No. 1.3e-08;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37  
:|||||  
Db 493 MFNNFTVSFWLRVPKVSASHLE 514

RESULT 12  
US-08-913-880C-13  
; Sequence 13, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913,880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 13  
; LENGTH: 865  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 451 to 1315 of SEQ ID NO: 1  
US-08-913-880C-13

Query Match 57.0%; Score 114; DB 4; Length 865;  
Best Local Similarity 95.5%; Pred. No. 1.3e-08;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37  
:|||||  
Db 496 MFNNFTVSFWLRVPKVSASHLE 517

RESULT 13  
US-08-913-880C-12  
; Sequence 12, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913,880C

; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 12  
; LENGTH: 866  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 450 to 1315 of SEQ ID NO: 1  
US-08-913-880C-12

Query Match 57.0%; Score 114; DB 4; Length 866;  
Best Local Similarity 95.5%; Pred. No. 1.3e-08;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37  
:|||||  
Db 497 MFNNFTVSFWLRVPKVSASHLE 518

RESULT 14  
US-08-913-880C-11  
; Sequence 11, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913,880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 11  
; LENGTH: 874  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 442 to 1315 of SEQ ID NO: 1  
US-08-913-880C-11

Query Match 57.0%; Score 114; DB 4; Length 874;  
Best Local Similarity 95.5%; Pred. No. 1.4e-08;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37  
:|||||  
Db 505 MFNNFTVSFWLRVPKVSASHLE 526

RESULT 15  
US-08-913-880C-10  
; Sequence 10, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913,880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 10  
; LENGTH: 875  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 441 to 1315 of SEQ ID NO: 1  
US-08-913-880C-10

Query Match 57.0%; Score 114; DB 4; Length 875;  
Best Local Similarity 95.5%; Pred. No. 1.4e-08;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37

Best Local Similarity 95.5%; Pred. No. 6.1e-09; DB 1; Length 618;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFVLRVPKVSASHLE 37  
DB 83 MFNNFTVSFVLRVPKVSASHLE 104

## RESULT 7

US-08-668-381A-5  
; Sequence 5, Application US/08668381A  
; Patent No. 5780024  
; GENERAL INFORMATION:  
; APPLICANT: Brown, Robert H.  
; APPLICANT: Fishman, Paul S.  
; APPLICANT: Francis, Jonathan W.  
; APPLICANT: Hosler, Betsy A.  
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN  
; TITLE OF INVENTION: FRAGMENT C HYBRID PROTEIN  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/668,381A  
; FILING DATE: 21-JUN-1996  
; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/000,473  
; FILING DATE: 23-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,164  
; REFERENCE/DOCKET NUMBER: 00786/269001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 618 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

## US-08-668-381A-5

Query Match 57.0%; Score 114; DB 1; Length 618;  
Best Local Similarity 95.5%; Pred. No. 8.9e-09;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFVLRVPKVSASHLE 37  
DB 249 MFNNFTVSFVLRVPKVSASHLE 270

## RESULT 8

US-08-913-880C-17  
; Sequence 17, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; TITLE OF INVENTION: VACCINE  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913,880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 15  
; LENGTH: 860  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 456 to 1315 of SEQ ID NO: 1  
; US-08-913-880C-15

Query Match 57.0%; Score 114; DB 4; Length 860;  
Best Local Similarity 95.5%; Pred. No. 1.3e-08;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

; CURRENT APPLICATION NUMBER: US/08/913,880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 17  
; LENGTH: 853  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 463 to 1315 of SEQ ID NO: 1  
; US-08-913-880C-17

Query Match 57.0%; Score 114; DB 4; Length 853;  
Best Local Similarity 95.5%; Pred. No. 1.3e-08;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFVLRVPKVSASHLE 37  
DB 484 MFNNFTVSFVLRVPKVSASHLE 505

## RESULT 9

US-08-913-880C-16  
; Sequence 16, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; TITLE OF INVENTION: VACCINE  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913,880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 16  
; LENGTH: 858  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 458 to 1315 of SEQ ID NO: 1  
; US-08-913-880C-16

Query Match 57.0%; Score 114; DB 4; Length 858;  
Best Local Similarity 95.5%; Pred. No. 1.3e-08;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFVLRVPKVSASHLE 37  
DB 489 MFNNFTVSFVLRVPKVSASHLE 510

## RESULT 10

US-08-913-880C-15  
; Sequence 15, Application US/08913880C  
; Patent No. 6372225  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Morihiro  
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
; TITLE OF INVENTION: VACCINE  
; FILE REFERENCE: 216-380P  
; CURRENT APPLICATION NUMBER: US/08/913,880C  
; CURRENT FILING DATE: 1997-11-24  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 15  
; LENGTH: 860  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
; FEATURE:  
; OTHER INFORMATION: Amino Acids 456 to 1315 of SEQ ID NO: 1  
; US-08-913-880C-15

Query Match 57.0%; Score 114; DB 4; Length 860;  
Best Local Similarity 95.5%; Pred. No. 1.3e-08;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

APPLICATION NUMBER: PCT/GB90/00943  
FILING DATE: 20-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8914122.0  
FILING DATE: 20 June 1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Mary J. Wilson  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 117-134  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-110-786A-8

Query Match 57.0%; Score 114; DB 1; Length 452;  
Best Local Similarity 95.5%; Pred. No. 6.1e-09;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNFTVSFWLRPKVSASHLE 37  
:|||||  
Db 83 MFNFTVSFWLRPKVSASHLE 104

## RESULT 5

US-08-280-228-2  
Sequence 2, Application US/08280228  
Patent No. 5571694  
GENERAL INFORMATION:  
APPLICANT: Makoff Dr, Andrew J  
APPLICANT: Romanos Dr, Michael A  
APPLICANT: Clare Dr, Jeffrey J  
APPLICANT: Fairweather Dr, Neil F  
TITLE OF INVENTION: VACCINES  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHUYE P.C.  
STREET: 1100 No. 5571694th Glebe Road  
CITY: Arlington,  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/280,228  
FILING DATE: 25-JUL-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/618,312  
FILING DATE: 27-NOV-1990  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8926832.0  
FILING DATE: 28-NOV-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9006097.1  
FILING DATE: 17-MAR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilson, Mary J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 117-163  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-280-228-2

Query Match 57.0%; Score 114; DB 1; Length 452;  
Best Local Similarity 95.5%; Pred. No. 6.1e-09;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNFTVSFWLRPKVSASHLE 37  
:|||||  
Db 83 MFNFTVSFWLRPKVSASHLE 104

## RESULT 6

US-08-280-228-4  
Sequence 4, Application US/08280228  
Patent No. 5571694  
GENERAL INFORMATION:  
APPLICANT: Makoff Dr, Andrew J  
APPLICANT: Romanos Dr, Michael A  
APPLICANT: Clare Dr, Jeffrey J  
APPLICANT: Fairweather Dr, Neil F  
TITLE OF INVENTION: VACCINES  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHUYE P.C.  
STREET: 1100 No. 5571694th Glebe Road  
CITY: Arlington,  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/280,228  
FILING DATE: 25-JUL-1994  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/618,312  
FILING DATE: 27-NOV-1990  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8926832.0  
FILING DATE: 28-NOV-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9006097.1  
FILING DATE: 17-MAR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilson, Mary J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 117-163  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-280-228-4

Query Match

57.0%; Score 114; DB 1; Length 452;

## RESULT 2

US-07-618-312A-2  
; Sequence 2, Application US/07618312A  
; Patent No. 5389540  
; GENERAL INFORMATION:  
; APPLICANT: Makoff Dr, Andrew J  
; APPLICANT: Romanos Dr, Michael A  
; APPLICANT: Clare Dr, Jeffrey J  
; APPLICANT: Fairweather Dr, Neil F  
; TITLE OF INVENTION: VACCINES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: 14th Floor  
; STREET: 2200 Clarendon Boulevard,  
; CITY: Arlington,  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/618,312A  
; FILING DATE: 19910516  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8926832.0  
; FILING DATE: 28-NOV-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crawford Mr, Arthur R  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 510-51  
; TELEPHONE: 0101 703 8750400  
; TELEFAX: 0101 703 5253468  
; TELEX: 200797 NIXN UR  
; INFORMATION FOR SEQ ID NO: 2:  
; LENGTH: 452 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-618-312A-2

Query Match 57.0%; Score 114; DB 1; Length 452;  
Best Local Similarity 95.5%; Pred. No. 6.1e-09;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37  
Db 83 MFNNFTVSFWLRVPKVSASHLE 104

## RESULT 3

US-07-618-312A-4  
; Sequence 4, Application US/07618312A  
; Patent No. 5389540  
; GENERAL INFORMATION:  
; APPLICANT: Makoff Dr, Andrew J  
; APPLICANT: Romanos Dr, Michael A  
; APPLICANT: Clare Dr, Jeffrey J  
; APPLICANT: Fairweather Dr, Neil F  
; TITLE OF INVENTION: VACCINES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: 14th Floor  
; STREET: 2200 Clarendon Boulevard,  
; CITY: Arlington,  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201

CITY: Arlington,  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/618,312A  
FILING DATE: 19910516  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8926832.0  
FILING DATE: 28-NOV-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9006097.1  
FILING DATE: 17-MAR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford Mr, Arthur R  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 510-51  
TELEPHONE: 0101 703 8750400  
TELEFAX: 0101 703 5253468  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-618-312A-4

Query Match 57.0%; Score 114; DB 1; Length 452;  
Best Local Similarity 95.5%; Pred. No. 6.1e-09;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37  
Db 83 MFNNFTVSFWLRVPKVSASHLE 104

## RESULT 4

US-08-110-786A-8  
; Sequence 8, Application US/08110786A  
; Patent No. 5443966  
; GENERAL INFORMATION:  
; APPLICANT: FAIRWEATHER, Neil Fraser  
; APPLICANT: MAKOFF, Andrew Joseph  
; TITLE OF INVENTION: Expression of tetanus toxin fragment C  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye P.C.  
; STREET: 1100 No. 5443966th Giebe Road  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/110,786A  
; FILING DATE: 23-AUG-1993 1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/777,337  
; FILING DATE: 29-NOV-1991  
; PRIOR APPLICATION DATA:

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:46:20 ; Search time 10.9668 Seconds  
(without alignments)  
142.750 Million cell updates/sec

Title: US-09-848-834A-14

Perfect score: 200

Sequence: 1 XHWSYGLRPGSSGSLFNFTVFWLRVVKVSASHLE 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_Patents\_AA:\*
- 1: /cgn2.6/ptodata/1/1aa/5A\_COMB.pep:\*
  - 2: /cgn2.6/ptodata/1/1aa/5B\_COMB.pep:\*
  - 3: /cgn2.6/ptodata/1/1aa/5A\_COMB.pep:\*
  - 4: /cgn2.6/ptodata/1/1aa/5B\_COMB.pep:\*
  - 5: /cgn2.6/ptodata/1/1aa/5A\_COMB.pep:\*
  - 6: /cgn2.6/ptodata/1/1aa/5B\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	114	57.0	31	5	PCT-US93-11703-64
2	114	57.0	452	1	US-07-618-312A-2
3	114	57.0	452	1	US-07-618-312A-2
4	114	57.0	452	1	US-08-110-786A-8
5	114	57.0	452	1	US-08-280-228-2
6	114	57.0	452	1	US-08-280-228-4
7	114	57.0	618	1	US-08-668-381A-5
8	114	57.0	853	4	US-08-913-880C-17
9	114	57.0	858	4	US-08-913-880C-16
10	114	57.0	860	4	US-08-913-880C-15
11	114	57.0	862	4	US-08-913-880C-14
12	114	57.0	865	4	US-08-913-880C-13
13	114	57.0	866	4	US-08-913-880C-12
14	114	57.0	874	4	US-08-913-880C-11
15	114	57.0	875	4	US-08-913-880C-10
16	114	57.0	1315	4	US-08-913-880C-1
17	112	56.0	21	1	US-07-610-525-1
18	112	56.0	21	2	US-08-661-052-12
19	112	56.0	21	2	US-08-460-502-8
20	112	56.0	21	2	US-08-724-774B-5
21	112	56.0	21	3	US-09-089-595-5
22	112	56.0	21	3	US-09-382-855-5
23	112	56.0	21	3	US-09-183-714B-5
24	112	56.0	21	3	US-09-188-082-12
25	112	56.0	21	3	US-09-171-969-10
26	112	56.0	21	4	US-09-364-088-12
27	112	56.0	21	4	US-09-642-281-5

28	112	56.0	21	4	US-09-102-716-12	Sequence 12, Appl
29	112	56.0	21	4	US-08-432-483A-3	Sequence 3, Appl
30	112	56.0	21	4	US-09-148-711A-8	Sequence 8, Appl
31	112	56.0	21	4	US-09-589-717-5	Sequence 5, Appl
32	112	56.0	21	4	US-08-945-289-3	Sequence 3, Appl
33	112	56.0	21	5	PCT-US93-11703-66	Sequence 66, Appl
34	112	56.0	32	1	US-08-446-692-14	Sequence 14, Appl
35	112	56.0	32	2	US-08-488-351A-14	Sequence 14, Appl
36	107	53.5	22	1	US-08-446-692-5	Sequence 41, Appl
37	107	53.5	22	2	US-08-488-351A-5	Sequence 5, Appl
38	107	53.5	22	3	US-09-100-409A-41	Sequence 41, Appl
39	107	53.5	22	5	PCT-US95-13841-8	Sequence 8, Appl
40	100	50.0	19	1	US-07-610-525-2	Sequence 2, Appl
41	94.5	47.2	20	2	US-08-319-704-11	Sequence 11, Appl
42	88	44.0	17	1	US-07-610-525-3	Sequence 3, Appl
43	87.5	43.8	22	2	US-08-817-933A-8	Sequence 8, Appl
44	79	39.5	14	1	US-07-610-525-6	Sequence 6, Appl
45	79	39.5	14	1	US-08-787-547-43	Sequence 43, Appl

## ALIGNMENTS

RESULT 1  
PCT-US93-11703-64  
; Sequence 64, Application PC/TUS93111703  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Mimotopes Pty. Ltd.  
; TITLE OF INVENTION: T-Cell Epitopes  
; NUMBER OF SEQUENCES: 75  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Grant D. Green  
; STREET: 4560 Horton St.  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/11703  
; FILING DATE: 28-DEC-1993  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/984,852  
; FILING DATE: 02-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Grant D.  
; REGISTRATION NUMBER: 31,259  
; REFERENCE/DOCKET NUMBER: 0222.101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 510-601-2706  
; TELEFAX: 510-655-3542  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; PCT-US93-11703-64

Query Match 57.0%; Score 114; DB 5; Length 31;  
Best Local Similarity 95.5%; Pred. No. 2.4e-10;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 16 LFNNFTVSWLRVVKVSASHLE 37  
:|||||  
DB 6 MFNNFTVSWLRVVKVSASHLE 27



```

XX WO200065058-A1.
XX PD 02-NOV-2000.
XX PF 19-APR-2000; 2000WO-DK00205.
XX PR 23-APR-1999; 99DK-0000552.
XX PR 06-MAY-1999; 99US-0132811.
XX PA (MEBI-) M & E BIOTECH AS.
XX PI Klysner S;
XX DR WPI; 2000-672791/65.
XX DR N-PSDB; AAC68868.
XX PT Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -
XX PS
XX PS Example 7; Page 141; 172pp; English.
XX CC The present invention is concerned with methods of treating asthma,
XX CC eosinophilia, allergic rhinitis and other allergic diseases. These
XX CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
XX CC proteins and their coding sequences to down-regulate IL-5 activity and
XX CC thus reduce eosinophil numbers. The allergic diseases may be treated
XX CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
XX CC it is possible that they may be used in the treatment of cancer and
XX CC helminthic infections.
XX SQ Sequence 124 AA;
XX
Query Match 58.5%; Score 117; DB 21; Length 124;
Best Local Similarity 95.7%; Pred. NO. 3e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 15 SLFNNFTVSFWLRVPKVSASHLE 37
DB 30 TLFNFTVSFWLRVPKVSASHLE 52

```

Search completed: September 24, 2003, 17:40:06  
Job time : 42.8172 secs

RESULT 14	
AAB45496	
ID	AAB45496 standard; Protein; 124 AA.
XX	
AC	AAB45496;
XX	
DT	26-FEB-2001 (first entry)
XX	
DE	Modified human interleukin-5 SEQ ID NO: 8.
XX	
KW	Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW	cancer; eosinophilia; vaccine; allergic rhinitis.
OS	Homo sapiens.
OS	Clostridium tetani.
XX	
PN	WO2000065058-A1.
XX	
PD	02-NOV-2000.
XX	
Pf	19-APR-2000; 2000WO-DK00205.
XX	
PR	23-APR-1999; 99DK-0000552.
PR	06-MAY-1999; 99US-0132811.
XX	
PA	(WEBI-) M & E BIOTECH AS.
XX	
PI	Klysner S;
XX	
DR	WPI; 2000-672791/65.
XX	
PT	Down-regulating interleukin 5 (IL-5) activity in humans by
PT	administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT	prophylaxis or amelioration of asthma or other chronic allergic
PT	conditions -
XX	
PS	Example 7; Page 124; 172pp; English.
XX	
CC	The present invention is concerned with methods of treating asthma,
CC	eosinophila, allergic rhinitis and other allergic diseases. These
CC	involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC	proteins and their coding sequences to down-regulate IL-5 activity and
CC	thus reduce eosinophil numbers. The allergic diseases may be treated
CC	using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC	it is possible that they may be used in the treatment of cancer and
CC	helminthic infections.
XX	
SQ	Sequence 124 AA;
Query Match	58.5%; Score 117; DB 21; Length 124;
Best Local Similarity	95.7%; Pred. No. 3e-09;
Matches	22; Conservative 1; Mismatches 0; Indels 0; Gaps
QY	15 SLFNNFTVSFWLVRPKVSASHLE 37
	:
Db	30 TLFNNTVSFWLVRPKVSASHLE 52
RESULT 15	
AAB45515	
ID	AAB45515 standard; Protein; 124 AA.
XX	
AC	AAB45515;
XX	
DT	26-FEB-2001 (first entry)
XX	
DE	Modified human interleukin-5 SEQ ID NO: 30.
XX	
KW	Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW	cancer; eosinophilia; vaccine; allergic rhinitis.
XX	
OS	Homo sapiens.
OS	Clostridium tetani.







Db 118 PGSTAPPANGVTSAPDTRFNNFTVSFWLRVPKVSASHLE 156

# RESULT 5

AAB20151  
ID AAB20151 standard; Protein; 109 AA.

AC AAB20151;

DT 30-APR-2001 (first entry)

XX Growth differentiation factor 8 AutoVac construct GDF-8 P30-3B.  
XX  
XX Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;  
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;  
KW cardiant; human; mutant; mutein.  
XX

OS Chimeric - Homo sapiens.  
OS Chimeric - Clostridium tetani.  
OS Synthetic.

XX Key Location/Qualifiers  
FH Region 1..83

FT /note= "identical to residues 267-349 of human  
FT GDF-8"

FT Region 84..104

FT /note= "tetanus toxoid P2 epitope"  
FT Region 105..109

FT /note= "identical to residues 371-375 of human  
FT GDF-8"

FT Misc-difference 73

FT /note= "Cys-73 may be substituted by Ser to avoid  
FT disulfide bond formation"

FT Misc-difference 90..91

FT /note= "optionally replaced by Glu-Gly"

PN WO200105820-A2.

PD 25-JAN-2001.

XX 20-JUL-2000; 2000WO-DK00413.

XX 20-JUL-1999; 99DK-0001014.  
XX 26-JUL-1999; 99US-0145275.

PA (MEBI-) M & E BIOTECH AS.

PI Halkier T, Mouritsen S, Klysner S;

XX WPI; 2001-112680/12.

XX Increasing the muscle mass of animals used in meat production by down  
PT regulating growth differentiation factor 8 (GDF-8) activity in the  
PT animal through induction of anti-GDF-8 antibody production

PS Example 1; Page 104; 110pp; English.

XX The present sequence is that of AutoVac construct GDF-8 P30-3B,  
CC comprising the 109 C-terminal amino acid residues of human  
CC growth differentiation factor 8 (GDF-8) in which residues 84-104  
CC are replaced by the promiscuous tetanus toxin T-cell epitope P30  
CC (see AAB20144). It is an object of the invention to produce a  
CC recombinant therapeutic vaccine that is capable of effecting  
CC down-regulation of GDF-8 in order to increase the muscle growth  
CC rate of farm animals. The vaccines (see AAB20145-53) are capable  
CC of breaking autotolerance against autologous GDF-8. They comprise  
CC the C-terminal portion of human GDF-8 in which a portion of the  
CC native sequence is replaced by a T-cell epitope such as P30, with  
CC minimal disturbance of the authentic 3-dimensional structure of  
CC the protein. Nucleic acids encoding the GDF-8 variants can be used  
CC for genetic immunisation of the animals. Down-regulation of GDF-8  
CC activity can increase muscle mass by up to at least 45% in cattle,  
CC pigs and poultry used for meat production, reducing the need for

CC antibiotic feed-additives. Anti-GDF8 vaccines can be used to  
CC treat human diseases such as cancer cachexia where muscle atrophy is  
CC pronounced and for patients suffering from acute and chronic heart  
CC failure.

XX Sequence 109 AA;

XX Query Match 59.8%; Score 119.5; DB 22; Length 109;  
XX Best Local Similarity 67.6%; Pred. NO. 1.1e-09;  
XX Matches 25; Conservative 1; Mismatches 2; Indels 9; Gaps 1;

QY 10 GSSGPSL-----FNNFTVSFWLRVPKVSASHLE 37

DB 68 GSAGPCCTPTKMSPIFNFTVSFWLRVPKVSASHLE 104

## RESULT 6

AY92644

ID AY92644 standard; Protein; 750 AA.

XX AC AY92644;

DT 10-AUG-2000 (first entry)

DE Mutant human prostate specific membrane antigen construct, hPSM6.3.

XX Prostate specific membrane antigen; immunogenized construct; mutant;  
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;  
KW prostate cancer; cell-associated peptide antigen; foreign epitope.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Peptide 448..462

FT /label= P2

FT /note= "foreign epitope"

FT Peptide 210..230

FT /label= P30

FT /note= "foreign epitope"

XX WO200020027-A2.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-DK00525.

XX 05-OCT-1998; 98DK-0001261.

XX 20-OCT-1998; 98US-0105011.

XX (MEBI-) M & E BIOTECH AS.

XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;

XX Gautam A, Birk P, Karlsson G;

XX WPI; 2000-349917/30.

XX Inducing immune responses to weakly immunogenic, tumor associated  
PT peptide antigens for the treatment of breast and prostate cancer  
PT  
XX Example 1; Page -; 220pp; English.

XX AY92627-49 are mutant immunogenized human prostate specific membrane  
CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or  
CC P30). The immunogenic analogues of PSM can be used in the claimed method  
CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,  
CC antibody binding regions and cysteine residues involved in disulfide  
CC bonds are preserved in the immunogenized forms. The method is used for  
CC inducing immune responses against weakly immunogenic cell-associated  
CC peptide antigens (PA) such as those associated with cancers  
CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),  
CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The  
CC method comprises effecting simultaneous presentation by antigen producing



XX PD 15-NOV-2001.  
 XX PF 04-MAY-2001; 2001WO-US14363.  
 XX PR 05-MAY-2000; 2000US-202328P.  
 XX PA (APHT-) APHTON CORP.  
 XX PI Grimes S, Michaeli D, Stevens VC;  
 XX DR WPI; 2002-049440/06.  
 XX PT Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue -  
 XX PS Claim 11; Page 9; 43pp; English.  
 XX CC The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX SQ Sequence 37 AA;  
 Query Match 99.5%; Score 199; DB 23; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 6e-22;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 HWSYGLRPGSSGSLFNFTVSFWLRVPKVSASHLE 37  
 Db 2 HWSYGLRPGSSGSLFNFTVSFWLRVPKVSASHLE 37  
 RESULT 2  
 AAU11429  
 ID AAU11429 standard; peptide; 50 AA.  
 AC AAU11429;  
 DT 12-MAR-2002 (first entry)  
 DE Synthetic immunogen peptide 10.  
 KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX Chimeric - Clostridium tetani.  
 OS Chimeric - Mammalia.  
 OS Synthetic.  
 XX Key Location/Qualifiers  
 FH Misc-difference 1  
 FT /label= OTHER  
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"  
 FT 1..10  
 FT Peptide  
 FT /note= "Gonadotropin releasing hormone epitope  
 FT (1..10 aa)"  
 FT 11..16  
 FT Peptide  
 FT /note= "Spacer peptide"

FT Peptide 17..37 "Tetanus toxoid (947-967 aa)"  
 FT /note= 38..41  
 FT Peptide /note= "Spacer peptide"  
 FT 42..50  
 FT Peptide /note= "Gonadotropin releasing hormone epitope  
 FT (2-10 aa)"  
 FT Modified-site 50  
 FT /note= "Amidated glycine or glycynamide"  
 XX WO200185763-A2.  
 XX 15-NOV-2001.  
 PD 04-MAY-2001; 2001WO-US14363.  
 PF 05-MAY-2000; 2000US-202328P.  
 PR (APHT-) APHTON CORP.  
 PA Grimes S, Michaeli D, Stevens VC;  
 XX WPI; 2002-049440/06.  
 XX Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue -  
 XX PS Claim 11; Page 11; 43pp; English.  
 XX CC The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX SQ Sequence 50 AA;  
 Query Match 99.5%; Score 199; DB 23; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-22;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 HWSYGLRPGSSGSLFNFTVSFWLRVPKVSASHLE 37  
 Db 2 HWSYGLRPGSSGSLFNFTVSFWLRVPKVSASHLE 37  
 RESULT 3  
 AAB20150  
 ID AAB20150 standard; Protein; 109 AA.  
 AC AAB20150;  
 XX 30-APR-2001 (first entry)  
 DT Growth differentiation factor 8 AutoVac construct GDF-8 P30-3A.  
 DE Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;  
 KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;  
 KW cardiant; human; mutant; mutein.  
 XX Chimeric - Homo sapiens.  
 OS Chimeric - Clostridium tetani.  
 OS Synthetic.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:58:36 ; Search time 41.8172 Seconds  
(without alignments)  
140.442 Million cell updates/sec

Title: US-09-848-834A-14

Perfect score: 200

Sequence: 1 XHWSYGLRPGSGSPFNFTVSWFLRPVKVASHLE 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

```
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	199	99.5	37	AAU11425	Synthetic immunogen
2	199	99.5	50	AAU11429	Synthetic immunogen
3	123	61.5	109	AAAB20150	Growth differentia
4	121	60.5	216	AAAY92665	MUC-1 analogue con
5	119.5	59.8	109	AAAB20151	Growth differentia
6	119.5	59.8	750	AAAY92644	Mutant human prost
7	119	59.5	122	AAAB45507	Modified murine int
8	119	59.5	122	AAAB45524	Modified murine in
9	118.5	59.2	158	AAW81332	TNF30-1, a TNF- $\alpha$ p

10	118.5	59.2	158	23	ABB07282	Human TNF- $\alpha$ an
11	118	59.0	109	22	AAAB20149	Growth differentia
12	118	59.0	703	21	AAAY92662	Mutant murine PSM
13	118	59.0	761	21	AAAY92660	Mutant murine pros
14	117	58.5	124	21	AAAB45496	Modified human int
15	117	58.5	124	21	AAAB45515	Modified human int
16	117	58.5	160	22	AAAB20153	Growth differentia
17	116	58.0	31	21	AAAY92653	PSMpep010 - P30 in
18	116	58.0	43	22	AAAB46177	Tetanus toxoid 830
19	116	58.0	43	22	AAAB49076	Anyloid beta/tetan
20	116	58.0	72	22	AAAB46190	Tetanus toxoid epi
21	116	58.0	109	22	AAAB20148	Growth differentia
22	116	58.0	136	22	AAAB49089	Anyloid beta tetan
23	116	58.0	145	21	AAAB45530	Modified murine in
24	116	58.0	147	21	AAAB45522	Modified murine int
25	116	58.0	158	19	AAW81334	TNF30-3, a TNF- $\alpha$ p
26	116	58.0	158	23	ABB07274	Human TNF- $\alpha$ an
27	116	58.0	188	21	AAAY84423	An osteoprotegerin
28	116	58.0	254	22	AAAB20152	Growth differentia
29	116	58.0	750	21	AAAY92627	Mutant human prost
30	116	58.0	750	21	AAAY92628	Mutant human prost
31	116	58.0	750	21	AAAY92629	Mutant human prost
32	116	58.0	750	21	AAAY92630	Mutant human prost
33	116	58.0	750	21	AAAY92631	Mutant human prost
34	116	58.0	750	21	AAAY92637	Mutant human prost
35	116	58.0	750	21	AAAY92638	Mutant human prost
36	116	58.0	750	21	AAAY92639	Mutant human prost
37	116	58.0	750	21	AAAY92642	Mutant human prost
38	115	57.5	693	21	AAAY92647	Mutant human PSM a
39	115	57.5	693	21	AAAY92648	Mutant human PSM a
40	115	57.5	750	21	AAAY92645	Mutant human prost
41	115	57.5	750	21	AAAY92646	Mutant human prost
42	114	57.0	130	21	AAAB45497	Modified human int
43	114	57.0	130	21	AAAB45516	Modified human int
44	114	57.0	452	12	AAAB2471	Tetanus toxin frag
45	114	57.0	453	22	AAAB31427	Amino acid sequenc

#### ALIGNMENTS

RESULT 1

AAU11425  
ID AAU11425 standard; peptide; 37 AA.

XX AC AAU11425;

XX XX 12-MAR-2002 (first entry)

XX DE Synthetic immunogen peptide 6.

XX KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;

XX KW luteinising hormone releasing hormone; GnRH; contraceptive;

XX KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;

XX KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;

XX KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX OS Chimeric - Clostridium tetani.

OS OS Chimeric - Mammalia.

OS OS Synthetic.

FH Key Location/Qualifiers

FT Misc-difference 1

FT /label= OTHER

FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"

FT Peptide

FT /note= "Gonadotrophin releasing hormone epitope"

FT Peptide

FT /note= "Spacer peptide"

FT Peptide

FT /note= "Tetanus toxoid sequence (947-967 aa)"

XX WO200185763-A2.



Query Match 56.0%; Score 112; DB 12; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.4e-08;  
Matches 21: Conservative 0; Mismatches 0 Indels

Query Match	56.0%; Score
Best Local Similarity	100.0%; Pre
Matches	21; Conservative 0; M
Qy	17 FNNFTVSEWLRVPKVSASHLE 37
Db	1 FNNFTVSEWLRVPKVSASHLE 21

**QY**            17 FNNFTVSFWLVRPKVSASHLE 37  
                |||||

**D6**            1 FNNFTVSFWLVRPKVSASHLE 21  
                |||||

Search completed: September 24, 2003, 17:33:09  
Job time : 19.5512 secs

US 03 403 380 2 ; Sequence 2, Application US/09405986 ; Publication No. US20030157115A1

US 03 403 380 2 ; Sequence 2, Application US/09405986 ; Publication No. US20030157115A1

US 03 403 380 2 ; Sequence 2, Application US/09405986 ; Publication No. US20030157115A1

US 03 403 380 2 ; Sequence 2, Application US/09405986 ; Publication No. US20030157115A1

US 03 403 380 2 ; Sequence 2, Application US/09405986 ; Publication No. US20030157115A1

US 03 403 380 2 ; Sequence 2, Application US/09405986 ; Publication No. US20030157115A1

US 03 403 380 2 ; Sequence 2, Application US/09405986 ; Publication No. US20030157115A1

US 03 403 380 2 ; Sequence 2, Application US/09405986 ; Publication No. US20030157115A1

US 03 403 380 2 ; Sequence 2, Application US/09405986 ; Publication No. US20030157115A1

US 03 403 380 2 ; Sequence 2, Application US/09405986 ; Publication No. US20030157115A1

```
Query Match      56.0%; Score 112; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 17 FNNFTVSFWLRVPKVSASHLE 37  
 |||||  
 Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 15

US-10-223-809A-6  
; Sequence 6, Application US/10223809A  
; Publication No. US20030157117A1

US-10-223-809A-6  
; Sequence 6, Application US/10223809A  
; Publication No. US20030157117A1

```

: APPLICANT: rasmussen, Peter Birk et al.
: TITLE OF INVENTION: No. US20030157111A1 Method for Down-Regulation of Amyloid
:
: FILE REFERENCE: 674542-2008
: CURRENT APPLICATION NUMBER: US/10/223,809A*
: CURRENT FILING DATE: 2002-08-20
:

```

```

: APPLICANT: rasmussen, Peter Birk et al.
: TITLE OF INVENTION: No. US20030157111A1 Method for Down-Regulation of Amyloid
:
: FILE REFERENCE: 674542-2008
: CURRENT APPLICATION NUMBER: US/10/223,809A*
: CURRENT FILING DATE: 2002-08-20
:

```

```

: APPLICANT: rasmussen, Peter Birk et al.
: TITLE OF INVENTION: No. US20030157111A1 Method for Down-Regulation of Amyloid
:
: FILE REFERENCE: 674542-2008
: CURRENT APPLICATION NUMBER: US/10/223,809A*
: CURRENT FILING DATE: 2002-08-20
:

```

```

: APPLICANT: rasmussen, Peter Birk et al.
: TITLE OF INVENTION: No. US20030157111A1 Method for Down-Regulation of Amyloid
:
: FILE REFERENCE: 674542-2008
: CURRENT APPLICATION NUMBER: US/10/223,809A*
: CURRENT FILING DATE: 2002-08-20
:

```

```

: APPLICANT: rasmussen, Peter Birk et al.
: TITLE OF INVENTION: No. US20030157111A1 Method for Down-Regulation of Amyloid
:
: FILE REFERENCE: 674542-2008
: CURRENT APPLICATION NUMBER: US/10/223,809A*
: CURRENT FILING DATE: 2002-08-20
:

```

```

: APPLICANT: rasmussen, Peter Birk et al.
: TITLE OF INVENTION: No. US20030157111A1 Method for Down-Regulation of Amyloid
:
: FILE REFERENCE: 674542-2008
: CURRENT APPLICATION NUMBER: US/10/223,809A*
: CURRENT FILING DATE: 2002-08-20
:

```

```

: APPLICANT: rasmussen, Peter Birk et al.
: TITLE OF INVENTION: No. US20030157111A1 Method for Down-Regulation of Amyloid
:
: FILE REFERENCE: 674542-2008
: CURRENT APPLICATION NUMBER: US/10/223,809A*
: CURRENT FILING DATE: 2002-08-20
:

```

; FILE REFERENCE: 1581.0920000  
; CURRENT APPLICATION NUMBER: US/10/130.973A  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: PCT/GB00/04644  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: GB 9928530.6  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: GB 008658.7  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 1112  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-130-973A-16

Query Match 57.0%; Score 114; DB 12; Length 1112;  
Best Local Similarity 95.5%; Pred. No. 5.3e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37  
Db 741 MFNFTVSFWLRVPKVSASHLE 762

## RESULT 10

US-10-241-596-141  
; Sequence 141, Application US/10241596  
; Publication No. US20030166238A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbiological Research Authority  
; APPLICANT: The Speywood Laboratory Limited  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130003  
; CURRENT APPLICATION NUMBER: US/10/241.596  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: US 09/242,689  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: US 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB 9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB 9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 175  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 141  
; LENGTH: 1315  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-10-241-596-141

Query Match 57.0%; Score 114; DB 12; Length 1315;  
Best Local Similarity 95.5%; Pred. No. 6.4e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37  
Db 946 MFNFTVSFWLRVPKVSASHLE 967

## RESULT 11

US-09-943-548-3  
; Sequence 3, Application US/09943548  
; Patent No. US20020042364A1  
; GENERAL INFORMATION:

; APPLICANT: Ritterhaus, Charles W.  
; APPLICANT: Thomas, Lawrence J.  
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVI  
; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2  
; CURRENT APPLICATION NUMBER: US/09/943,548  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 08/432,483  
; PRIOR FILING DATE: 1995-05-01  
; PRIOR APPLICATION NUMBER: PCT/US96/06147  
; PRIOR FILING DATE: 1996-05-01  
; PRIOR APPLICATION NUMBER: 08/945,289  
; PRIOR FILING DATE: 1997-10-17  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: helper T cell epitope of tetanus toxin  
US-09-943-548-3

Query Match 56.0%; Score 112; DB 9; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.4e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37  
Db 1 FNNFTVSFWLRVPKVSASHLE 21

## RESULT 12

US-09-848-834A-4  
; Sequence 4, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Tetanus bacillus  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(21)  
; OTHER INFORMATION: Amino acid sequence 947-967 of tetanus  
; OTHER INFORMATION: Toxoid Precursor (Tentoxylisin)  
US-09-848-834A-4

Query Match 56.0%; Score 112; DB 9; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.4e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37  
Db 1 FNNFTVSFWLRVPKVSASHLE 21

## RESULT 13

US-09-785-215-6  
; Sequence 6, Application US/09785215  
; Publication No. US20020187157A1  
; GENERAL INFORMATION:  
; APPLICANT: JENSEN, Martin Roland et al.  
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID  
; FILE REFERENCE: 3631-0107P

```
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: PCT/GB00/04644
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: GB 9928530.6
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: GB 008658.7
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-130-973A-9
```

```
Query Match 57.0%; Score 114; DB 12; Length 665;
Best Local Similarity 95.5%; Pred. No. 3e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 294 MFNNFTVSFWLRVPKVSASHLE 315
:|||||
```

## RESULT 6

```
US-10-130-973A-3
; Sequence 3, Application US/10130973A
; Publication No. US20030147895A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford
; APPLICANT: Sutton, John
; APPLICANT: Silman, Nigel
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells
; FILE REFERENCE: 1581.0920000
; CURRENT APPLICATION NUMBER: US/10/130,973A
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: PCT/GB00/04644
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: GB 9928530.6
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: GB 008658.7
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-130-973A-3
```

```
Query Match 57.0%; Score 114; DB 12; Length 882;
Best Local Similarity 95.5%; Pred. No. 4.1e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 511 MFNNFTVSFWLRVPKVSASHLE 532
:|||||
```

## RESULT 7

```
US-10-130-973A-5
; Sequence 5, Application US/10130973A
; Publication No. US20030147895A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford
; APPLICANT: Sutton, John
; APPLICANT: Silman, Nigel
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells
; FILE REFERENCE: 1581.0920000
```

```
; CURRENT APPLICATION NUMBER: US/10/130,973A
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: PCT/GB00/04644
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: GB 9928530.6
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: GB 008658.7
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-130-973A-5
```

```
Query Match 57.0%; Score 114; DB 12; Length 907;
Best Local Similarity 95.5%; Pred. No. 4.3e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 536 MFNNFTVSFWLRVPKVSASHLE 557
:|||||
```

## RESULT 8

```
US-10-130-973A-17
; Sequence 17, Application US/10130973A
; Publication No. US20030147895A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford
; APPLICANT: Sutton, John
; APPLICANT: Silman, Nigel
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells
; FILE REFERENCE: 1581.0920000
; CURRENT APPLICATION NUMBER: US/10/130,973A
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: PCT/GB00/04644
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: GB 9928530.6
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: GB 008658.7
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 1052
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-130-973A-17
```

```
Query Match 57.0%; Score 114; DB 12; Length 1052;
Best Local Similarity 95.5%; Pred. No. 5e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 681 MFNNFTVSFWLRVPKVSASHLE 702
:|||||
```

## RESULT 9

```
US-10-130-973A-16
; Sequence 16, Application US/10130973A
; Publication No. US20030147895A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford
; APPLICANT: Sutton, John
; APPLICANT: Silman, Nigel
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells
```

Best Local Similarity 100.0%; Pred. No. 3e-20;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSGSLFNNFTVFWLRVPKVSASHLE 37  
|||||  
Db 2 HWSYGLRPGSGSLFNNFTVFWLRVPKVSASHLE 37

RESULT 2  
US-09-848-834A-18  
; Sequence 18, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 18  
; LENGTH: 50  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human GnRH linked by a spacer to amino acid sequence 947-967 of the Tetanus toxin precursor (Tetanus toxin precursor (Tetoxylisin) protein linked by a spacer to amino acid sequence 2-10 of human GnRH  
; NAME/KEY: MOD.RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: MOD.RES  
; LOCATION: (50)..(50)  
; OTHER INFORMATION: Amidated glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(37)  
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxin precursor (Tetanus toxin precursor (Tetoxylisin) protein linked by a spacer to amino acid sequence 2-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (38)..(41)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (42)..(50)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-18

Query Match 99.5%; Score 199; DB 9; Length 50;  
Best Local Similarity 100.0%; Pred. No. 4.1e-20;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSGSLFNNFTVFWLRVPKVSASHLE 37  
|||||  
Db 2 HWSYGLRPGSGSLFNNFTVFWLRVPKVSASHLE 37

RESULT 3  
US-09-816-467-2  
; Sequence 2, Application US/09816467  
; Publication No. US20030004121A1  
; GENERAL INFORMATION:  
; APPLICANT: COEN, LAURENT  
; APPLICANT: PINZOLAS, ROSARIO OSTA  
; APPLICANT: BRULET, PHILIPPE  
; TITLE OF INVENTION: HYBRID PROTEINS THAT MIGRATE RETROGRADELY AND

; TITLE OF INVENTION: TRANSYNAPTICALLY INTO THE CNS  
; FILE REFERENCE: 03495.0174-01000  
; CURRENT APPLICATION NUMBER: US/09/816,467  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: 60/055,615  
; PRIOR FILING DATE: 1997-08-14  
; PRIOR APPLICATION NUMBER: 60/065,236  
; PRIOR FILING DATE: 1997-11-13  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 463  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-09-816-467-2

Query Match 57.0%; Score 114; DB 11; Length 463;  
Best Local Similarity 95.5%; Pred. No. 2.1e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVFWLRVPKVSASHLE 37  
:|||||  
Db 94 MFNNFTVFWLRVPKVSASHLE 115

RESULT 4  
US-10-130-973A-11  
; Sequence 11, Application US/10130973A  
; Publication No. US20030147895A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford  
; APPLICANT: Sutton, John  
; APPLICANT: Silman, Nigel  
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cell  
; FILE REFERENCE: 1581.0920000  
; CURRENT APPLICATION NUMBER: US/10/130,973A  
; CURRENT FILING DATE: 2002-10-21  
; PRIOR APPLICATION NUMBER: PCT/GB00/04644  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: GB 9928530.6  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: GB 008658.7  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 605  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-130-973A-11

Query Match 57.0%; Score 114; DB 12; Length 605;  
Best Local Similarity 95.5%; Pred. No. 2.8e-07;  
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVFWLRVPKVSASHLE 37  
:|||||  
Db 234 MFNNFTVFWLRVPKVSASHLE 255

RESULT 5  
US-10-130-973A-9  
; Sequence 9, Application US/10130973A  
; Publication No. US20030147895A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford  
; APPLICANT: Sutton, John  
; APPLICANT: Silman, Nigel  
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cell  
; FILE REFERENCE: 1581.0920000  
; CURRENT APPLICATION NUMBER: US/10/130,973A

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:57:10 ; Search time 18.5512 seconds  
(without alignments)  
301.778 Million cell updates/sec

Title: US-09-848-834A-14

Perfect score: 200

Sequence: 1 XHWSYGLRPGSGPSLFNFTVSFWLRVPRKVSASHLE 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	199	99.5	37	9 US-09-848-834A-14	Sequence 14, Appl
2	199	99.5	50	9 US-09-848-834A-18	Sequence 18, Appl
3	114	57.0	463	11 US-09-816-467-2	Sequence 2, Appl
4	114	57.0	605	12 US-10-130-973A-11	Sequence 11, Appl
5	114	57.0	665	12 US-10-130-973A-9	Sequence 9, Appl
6	114	57.0	882	12 US-10-130-973A-3	Sequence 3, Appl
7	114	57.0	907	12 US-10-130-973A-5	Sequence 5, Appl
8	114	57.0	1052	12 US-10-130-973A-17	Sequence 17, Appl
9	114	57.0	1112	12 US-10-130-973A-16	Sequence 16, Appl
10	114	57.0	1315	12 US-10-241-596-141	Sequence 141, App
11	112	56.0	21	9 US-09-943-548-3	Sequence 3, Appl
12	112	56.0	21	9 US-09-848-834A-4	Sequence 4, Appl
13	112	56.0	21	9 US-09-785-215-6	Sequence 6, Appl
14	112	56.0	21	12 US-09-405-986-2	Sequence 2, Appl
15	112	56.0	21	12 US-10-223-809A-6	Sequence 6, Appl

16	112	56.0	21	12	US-10-261-208-5	Sequence 5, Appl
17	112	56.0	21	15	US-10-204-362-6	Sequence 6, Appl
18	112	56.0	21	15	US-10-339-522-3	Sequence 3, Appl
19	112	56.0	21	15	US-10-223-711-8	Sequence 8, Appl
20	112	56.0	34	9	US-09-848-834A-10	Sequence 10, Appl
21	107	53.5	22	11	US-09-865-294-4	Sequence 4, Appl
22	95.5	47.8	20	14	US-10-044-034-20	Sequence 20, Appl
23	88	44.0	22	11	US-09-747-802-12	Sequence 12, Appl
24	87	43.5	31	9	US-09-848-834A-15	Sequence 15, Appl
25	87	43.5	34	9	US-09-848-834A-13	Sequence 13, Appl
26	87	43.5	36	9	US-09-848-834A-16	Sequence 16, Appl
27	87	43.5	46	9	US-09-848-834A-19	Sequence 19, Appl
28	87	43.5	47	9	US-09-848-834A-17	Sequence 17, Appl
29	87	43.5	51	9	US-09-848-834A-20	Sequence 20, Appl
30	79	39.5	14	10	US-09-909-460-43	Sequence 43, Appl
31	79	39.5	15	10	US-09-909-460-44	Sequence 44, Appl
32	71	35.5	17	11	US-09-305-924-4	Sequence 4, Appl
33	67	33.5	449	11	US-09-910-186A-18	Sequence 18, Appl
34	65	32.5	425	9	US-09-288-326-9	Sequence 9, Appl
35	65	32.5	434	11	US-09-910-186A-4	Sequence 4, Appl
36	65	32.5	435	11	US-09-910-186A-6	Sequence 6, Appl
37	65	32.5	437	11	US-09-910-186A-2	Sequence 2, Appl
38	65	32.5	1295	10	US-09-726-949A-1	Sequence 1, Appl
39	64.5	32.2	49	9	US-09-019-010-4	Sequence 4, Appl
40	64.5	32.2	49	11	US-09-305-924-11	Sequence 11, Appl
41	64.5	32.2	695	11	US-09-305-924-13	Sequence 13, Appl
42	63	31.5	440	11	US-09-910-186A-8	Sequence 8, Appl
43	63	31.5	1169	12	US-10-241-596-20	Sequence 20, Appl
44	61	30.5	16	11	US-09-305-924-5	Sequence 5, Appl
45	61	30.5	144	8	US-08-981-087A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-848-834A-14  
; Sequence 14, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphcon Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 14  
; LENGTH: 37  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the  
; OTHER INFORMATION: nrh hormone linked by a spacer to amino acid sequence 947-967-  
; OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylisin)  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(37)  
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor  
; OTHER INFORMATION: (Tentoxylisin)  
US-09-848-834A-14

Query Match 99.5%; Score 199; DB 9; Length 37;



RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,  
 RA Phillips H.S., Nikolics K., Seeburg P.H.;  
 RT "A deletion truncating the gonadotropin-releasing hormone gene is  
 RT responsible for hypogonadism in the hpg mouse.";  
 RL Science 234:1366-1371(1986).  
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the GnRH family.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M14872; AAA37717.1; -;  
 DR PIR: A47578; RMSG.  
 DR MGD: MGI:95789; GnRH.  
 DR InterPro: IPR002012; GnRH.  
 DR InterPro: IPR004079; GonadoliberinI.  
 DR Pfam: PF00446; GnRH; 1.  
 DR PRINTS: PR01541; GONADOLIBERNI.  
 DR PROSITE: PS00473; GnRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Signal; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 21  
 FT CHAIN 22 90 PROGONADOLIBERIN I.  
 FT PEPTIDE 22 31 GONADOLIBERIN I.  
 FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.  
 FT ACT\_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.  
 FT MOD\_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;  
 Query Match 29.0%; Score 58; DB 1; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 HWSYGLRPG 10  
 Db 23 HWSYGLRPG 31  
 RESULT 15  
 ID GONL\_PIG STANDARD; PRT; 91 AA.  
 AC P49221;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonadoliberin I precursor (Contains: Gonadoliberin I (LH-RH I)  
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].  
 GN GNRHI OR GNRH.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Hypothalamus;  
 RA Weesner G.D., Matteri R.L., Becker B.A.;  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE OF 24-33.  
 RX MEDLINE=72114303; PubMed=4946067;  
 RA Baba Y., Matsuo H., Schally A.V.;  
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.

RT Confirmation of the proposed structure by conventional sequential  
 RT analyses";  
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).  
 RN [3]  
 RP SYNTHESIS OF GONADOLIBERIN.  
 RX MEDLINE=72065376; PubMed=4942736;  
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;  
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-  
 RT phase method.";  
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).  
 RN [4]  
 RP SYNTHESIS OF GONADOLIBERIN.  
 RX MEDLINE=72117544; PubMed=4946275;  
 RA Baba Y., Arimura A., Schally A.V.;  
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";  
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).  
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the GnRH family.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: L32864; AAA31066.1; -;  
 DR InterPro: IPR002012; GnRH.  
 DR InterPro: IPR004079; GonadoliberinI.  
 DR Pfam: PF00446; GnRH; 1.  
 DR PRINTS: PR01541; GONADOLIBERNI.  
 DR PROSITE: PS00473; GnRH; 1.  
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Signal; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 23  
 FT CHAIN 24 91 PROGONADOLIBERIN I.  
 FT PEPTIDE 24 33 GONADOLIBERIN I.  
 FT PEPTIDE 34 91 GnRH-ASSOCIATED PEPTIDE I.  
 FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.  
 FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;  
 Query Match 29.0%; Score 58; DB 1; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 0.25;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 HWSYGLRPG 10  
 Db 25 HWSYGLRPG 33  
 Search completed: September 24, 2003, 17:41:19  
 Job time : 6.96704 secs

```

AC 009163;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
GN GNRH1 OR GNRH OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the GNRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U91938; AAB51302.1; -.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadolibirinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBRN1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pyrrolidone carboxylic acid.
FT NON_TER 1 1
FT CHAIN 1 >63 PROGONADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >63 GNRH-ASSOCIATED PEPTIDE I (BY
FT SIMILARITY).
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY (BY SIMILARITY).
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
FT NON_TER 63 63
FT SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;

Query Match 29.0%; Score 58; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
Db 2 HWSYGLRPG 10
|||||
2 HWSYGLRPG 10

RESULT 13
GONI_XENLA
ID GONI_XENLA STANDARD; PRT; 89 AA.
AC P45656;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadolibirin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
DE (LH-RH) (Luliberin I).
OS Xenopus laevis (African clawed frog).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Forebrain;
RA MEDLINE=9418563; PubMed=8137750;
RA Hayes W.P., Gray S., Battey J.F.;
RA "The frog gonadotropin-releasing hormone-I (GnRH-I) gene has a
RA mammalian-like expression pattern and conserved domains in
RA GnRH-associated peptide, but brain onset is delayed until
RA metamorphosis."
RL Endocrinology 134:1835-1844(1994).
CC -1- FUNCTION: Stimulates the secretion of gonadotropins.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the GNRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L28040; AAA49728.1; -.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadolibirinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBRN1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23 PROGONADOLIBERIN I.
FT CHAIN 24 89 GONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 89 GONADOTROPIN-RELEASING HORMONE ASSOCIATED
FT PEPTIDE.
FT PEPTIDE 37 85 GNRH-ASSOCIATED PEPTIDE I (GAP).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT SEQUENCE 89 AA; 10246 MW; 6F4F36FBAE0D4284 CRC64;

Query Match 29.0%; Score 58; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33
|||||
25 HWSYGLRPG 33

RESULT 14
GONI_MOUSE
ID GONI_MOUSE STANDARD; PRT; 90 AA.
AC P13562;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE I].
GN GNRH1 OR GNRH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87069928; PubMed=3024317;

```

RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [6]  
 RP ERRATUM.  
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,  
 RA Lander E.S.;  
 RL Nat. Genet. 23:373-373(1999).  
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),  
 CC Lutrepulse or Lutrelief (Ferring Pharmaceuticals) and Relisorm  
 CC (Serono).  
 CC -!- SIMILARITY: Belongs to the GnRH family.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X01059; CAA25526.1; -;  
 CC EMBL; M12578; AAA35916.1; -;  
 CC EMBL; X15215; CAA33285.1; -;  
 CC PIR; S05308; RHUG.  
 CC Genew; HGNC:4419; GNRH1.  
 CC MIM; 152760; -;  
 CC  
 CC GO; GO:0005625; C:soluble fraction; TAS.  
 CC GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.  
 CC GO; GO:0007267; P:cell-cell signaling; TAS.  
 CC GO; GO:0007275; P:development; TAS.  
 CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
 CC GO; GO:0007165; P:signal transduction; TAS.  
 CC InterPro; IPR002012; GnRH.  
 CC InterPro; IPR004079; GonadoliberinI.  
 CC Pfam; PF00446; GnRH; 1.  
 CC PRINTS; PR01541; GONADOLIBRNI.  
 CC PROSITE; PS00473; GnRH; 1.  
 CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 CC Placenta; Pharmaceutical; Signal; Polymorphism;  
 CC Pyrrolidone carboxylic acid.  
 CC SIGNAL 1 23  
 FT CHAIN 24 92 PROGONADOLIBERIN I.  
 FT PEPTIDE 24 33 GONADOLIBERIN I.  
 FT PEPTIDE 37 92 GnRH-ASSOCIATED PEPTIDE I.  
 FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.  
 FT MOD\_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 FT VARIANT 16 16 W->S (IN dbSNP:6185).  
 FT /FTID-Var\_013943.  
 SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;  
 Query Match 29.5%; Score 59; DB 1; Length 92;  
 Best Local Similarity 52.4%; Pred. No. 0.18;  
 Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;  
 QY 2 HWSYGLRPGS--SGPSLFNPF 20  
 Db 25 HWSYGLRPGGKRDAENLIDSF 45  
 RESULT 11  
 GONI\_SHEEP STANDARD; PRT; 61 AA.  
 ID GONI\_SHEEP

AC Q28588;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)  
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
 DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]  
 DE (Fragment).  
 GN GNRH1 OR GnRH OR LHRH.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE OF 12-61 FROM N.A.  
 RA STRAIN=Western range; TISSUE=Hypothalamus;  
 RC Rodriguez R.E., Wise M.E.;  
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-10.  
 RX MEDLINE=72094314; PubMed=4550508;  
 RA Burkus R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,  
 RA Fellows R., Blackwell R., Vale W., Guillemin R.;  
 RT "Primary structure of the ovine hypothalamic luteinizing hormone-  
 RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass  
 RT spectrometry-decapeptide-Edman degradation).";  
 RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).  
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the GnRH family.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U02517; AAA03433.1; -;  
 CC InterPro; IPR002012; GnRH.  
 CC InterPro; IPR004079; GonadoliberinI.  
 CC Pfam; PF00446; GnRH; 1.  
 CC PRINTS; PR01541; GONADOLIBRNI.  
 CC PROSITE; PS00473; GnRH; 1.  
 CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 CC Placenta; Pyrrolidone carboxylic acid.  
 CC NON\_TER 1 1  
 FT CHAIN 1 >61 PROGONADOLIBERIN I.  
 FT PEPTIDE 1 10 GONADOLIBERIN I.  
 FT PEPTIDE 14 >61 GnRH-ASSOCIATED PEPTIDE I.  
 FT ACT\_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT ACTIVITY.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).  
 FT NON\_TER 61 61  
 SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;  
 Query Match 29.0%; Score 58; DB 1; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 0.16;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 HWSYGLRPG 10  
 Db 2 HWSYGLRPG 10  
 RESULT 12  
 GONI\_MESAU STANDARD; PRT; 63 AA.  
 ID GONI\_MESAU

[illegible]

Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 11 SSGPSLENNFTVSFWLRVPKVSASHLE 37  
Db 630 TEGVTFSKRFSVYWRVEKITRKHLE 656

RESULT 7  
BXA2\_CLOBO STANDARD; PRT; 1295 AA.  
AC Q45894; P77780;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (Bont/A)  
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-chain; Botulinum neurotoxin A, heavy-chain].  
GN BOTA OR BNA OR ATX.  
OS Clostridium botulinum.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.  
OX NCBI\_TaxID=1491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Kyoto-F;  
RX MEDLINE=94143603; PubMed=8310180;  
RA Willems A., East A.K., Lawson P.A., Collins M.D.;  
RT "Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A associated with infant botulism: comparison with other clostridial neurotoxins.";  
RL Res. Microbiol. 144:547-556(1993).  
RN [2]  
RP SEQUENCE OF 1-65 FROM N.A.  
RC STRAIN=Kyoto-F;  
RX MEDLINE=97016817; PubMed=8863443;  
RA East A.K., Bhandari M., Stacey J.N., Campbell K.D., Collins M.D.;  
RT "Organizational and phylogenetic interrelationships of genes encoding components of the botulinum toxin complex in proteolytic Clostridium botulinum types A, B, and F: evidence of chimeric sequences in the gene encoding the nontoxic nonhemagglutinin component.";  
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).  
CC -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin binds with high affinity to peripheral neuronal presynaptic membrane, is then internalized by receptor-mediated endocytosis. The C-terminus of the heavy chain (H) is responsible for the adherence of the toxin to the cell surface while the N-terminus mediates transport of the light chain from the endocytic vesicle to the cytosol. After translocation, the light chain (L) hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking neurotransmitter release. Inhibition of acetylcholine release results in flaccid paralysis, with frequent heart or respiratory failure (By similarity).  
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No detected action on small molecule substrates.  
CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a heavy chain (H) (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of botulinum neurotoxin: Types A, B, C1, D, E, F, and G.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X73423; CAA51824.1; -;  
CC EMBL; X87974; CAA61234.1; -;  
CC PIR; I40645; I40645.

DR HSP; P10845; 3BTA.  
DR MEROPS; M27.002; -;  
DR InterPro; IPR000395; Bontoxilysin.  
DR Pfam; PF01742; Peptidase\_M27; 1.  
DR ProDom; PD001963; Bontoxilysin; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; FALSE\_NEG.  
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.  
FT INIT\_MET 0  
FT CHAIN 1 447  
FT CHAIN 448 1295  
FT METAL 222 222  
FT ACT\_SITE 223 223  
FT METAL 226 226  
FT DISULFID 429 453  
FT DISULFID 1234 1279  
FT TRANSMEM 626 646  
FT TRANSMEM 655 675  
SQ SEQUENCE 1295 AA; 149279 MW; 5DA04A13D98D6372 CRC64;

Query Match 31.0%; Score 62; DB 1; Length 1295;  
Best Local Similarity 50.0%; Pred. No. 1.3;  
Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30  
Db 935 SMYENFTSFWIKIPK 950  
I::||:|||||

RESULT 8  
GONL\_TUPGB STANDARD; PRT; 92 AA.  
AC Q95335;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Gonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I) (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide II].  
GN GNRHI OR GNRH.  
OS Tupaiia glis belangeri (Common tree shrew).  
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.  
OX NCBI\_TaxID=9396;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Hypothalamus;  
RX MEDLINE=97079639; PubMed=8921350;  
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P., Fernald R.D.;  
RT "Characterization of two new preproGNRH mRNAs in the tree shrew: first direct evidence for mesencephalic GNRH gene expression in a placental mammal.";  
RL Gen. Comp. Endocrinol. 104:7-19(1996).  
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS: IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the GNRH family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U63326; AAB16837.1; -;  
DR InterPro; IPR002012; GnrH.  
DR Pfam; PF00446; GnrH; 1.  
DR PRINTS; PR01541; GONADOLIBRNI.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:46:20 ; Search time 10.076 Seconds  
(without alignments)  
142.750 Million cell updates/sec

Title: US-09-848-834A-10

Perfect score: 186

Sequence: 1 FNNFTVSPWLVRPKVSASHLEGPRLHMSYGLRXP 34

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	81.7	32	1	US-08-446-692-14
2	152	81.7	32	2	US-08-488-351A-14
3	112	60.2	21	1	US-07-610-525-1
4	112	60.2	21	2	US-08-661-052-12
5	112	60.2	21	2	US-08-460-502-8
6	112	60.2	21	2	US-08-724-774B-5
7	112	60.2	21	3	US-09-089-595-5
8	112	60.2	21	3	US-09-382-855-5
9	112	60.2	21	3	US-09-183-714B-5
10	112	60.2	21	3	US-09-188-082-12
11	112	60.2	21	3	US-09-171-969-10
12	112	60.2	21	4	US-09-364-088-12
13	112	60.2	21	4	US-09-642-281-5
14	112	60.2	21	4	US-09-102-716-12
15	112	60.2	21	4	US-08-432-483A-3
16	112	60.2	21	4	US-09-148-711A-8
17	112	60.2	21	4	US-09-589-717-5
18	112	60.2	21	4	US-08-945-289-3
19	112	60.2	21	5	PCR-US93-11703-66
20	112	60.2	452	1	US-07-618-312A-2
21	112	60.2	452	1	US-07-618-312A-4
22	112	60.2	452	1	US-08-110-786A-8
23	112	60.2	452	1	US-08-280-228-2
24	112	60.2	452	1	US-08-280-228-4
25	112	60.2	618	1	US-08-668-381A-5
26	112	60.2	853	4	US-08-913-880C-17
27	112	60.2	853	4	US-08-913-880C-17

28	112	60.2	858	4	US-08-913-880C-16	Sequence 16, Appl
29	112	60.2	860	4	US-08-913-880C-15	Sequence 15, Appl
30	112	60.2	862	4	US-08-913-880C-14	Sequence 14, Appl
31	112	60.2	865	4	US-08-913-880C-13	Sequence 13, Appl
32	112	60.2	866	4	US-08-913-880C-12	Sequence 12, Appl
33	112	60.2	874	4	US-08-913-880C-11	Sequence 11, Appl
34	112	60.2	875	4	US-08-913-880C-10	Sequence 10, Appl
35	112	60.2	1315	4	US-08-913-880C-1	Sequence 1, Appl
36	107	57.5	22	1	US-08-446-692-5	Sequence 5, Appl
37	107	57.5	22	2	US-08-488-351A-5	Sequence 4, Appl
38	107	57.5	22	3	US-09-100-409A-41	Sequence 3, Appl
39	107	57.5	22	5	PCR-US93-13841-8	Sequence 2, Appl
40	100	53.8	19	1	US-07-610-525-2	Sequence 2, Appl
41	94.5	50.8	20	2	US-08-319-704-11	Sequence 11, Appl
42	88	47.3	17	1	US-07-610-525-3	Sequence 3, Appl
43	87.5	47.0	22	2	US-08-817-933A-8	Sequence 8, Appl
44	79	42.5	14	1	US-07-610-525-6	Sequence 6, Appl
45	79	42.5	14	1	US-08-787-547-43	Sequence 43, Appl

## ALIGNMENTS

RESULT 1  
US-08-446-692-14  
Sequence 14, Application US/08446692

Patent No. 5759551

GENERAL INFORMATION:

APPLICANT: Ladd, Anna

APPLICANT: Wang, Chang Yi

TITLE OF INVENTION: Immunogenic LHRH peptide constructs

TITLE OF INVENTION: And synthetic universal immune stimulators for vaccines

NUMBER OF SEQUENCES: 114

CORRESPONDENCE ADDRESS:

ADDRESSEE: Maria C.H. Lin

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Maria C.H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4146 US2

TELEPHONE: (212)415-8745

TELEFAX: (516)751-6849

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 32 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-446-692-14

Query Match 81.7%; Score 152; DB 1; Length 32;

Best local Similarity 87.9%; Pred. No. 1.6e-15;

Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 FNNFTVSPWLVRPKVSASHLEGPRLHMSYGLRXP 33  
|||||  
DB 3 FNNFTVSPWLVRPKVSASHLEGPRLHMSYGLRXP 31

RESULT 2  
US-08-488-351A-14  
; Sequence 14, Application US/08488351A  
; Patent No. 5843446  
; GENERAL INFORMATION:  
; APPLICANT: Ladd, Anna  
; APPLICANT: Wang, Chang Y1  
; APPLICANT: Zamb, Timothy  
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maria C.H. Lin  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: US  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488, 351A  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/446,692  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/229, 275  
; FILING DATE: 14-APR-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/057,166  
; FILING DATE: 27-APR-1992  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria C.H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)415-8745  
; TELEFAX: (516)751-6849  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-488-351A-14

Query Match 81.7%; Score 152; DB 2; Length 32;  
Best Local Similarity 87.9%; Pred. No. 1.6e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

OY 1 FNNFTVSPMLRVKPVKSASHLEPSLHMSYGLRP 33  
Db 3 FNNFTVSPMLRVKPVKSASHLE---HMSYGLRP 31

RESULT 3  
US-07-610-525-1  
; Sequence 1, Application US/07610525  
; Patent No. 5196512  
; GENERAL INFORMATION:  
; APPLICANT: BIANCHI Elisabetta  
; APPLICANT: PESSI Antonello  
; APPLICANT: CORRADIN Giampietro  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES USEFUL AS

; TITLE OF INVENTION: UNIVERSAL CARRIERS FOR THE PREPARATION OF IMMUNOGENIC CONJ  
; TITLE OF INVENTION: AND THEIR USE IN THE DEVELOPMENT OF SYNTHETIC VACCINES.  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHEA & GOULD  
; STREET: 1251 AVENUE OF THE AMERICAS  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: UNITED STATES  
; ZIP: 10020-1193  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/610,525  
; FILING DATE: 19901108  
; CLASSIFICATION: 424  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-827-3000  
; TELEFAX: 212-840-6702  
; TELEX: 423973  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acid residues  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: no  
; FRAGMENT TYPE: internal fragment  
US-07-610-525-1

Query Match 60.2%; Score 112; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNNFTVSPMLRVKPVKSASHLE 21  
Db 1 FNNFTVSPMLRVKPVKSASHLE 21

RESULT 4  
US-08-661-052-12  
; Sequence 12, Application US/08661052  
; Patent No. 5837243  
; GENERAL INFORMATION:  
; APPLICANT: Yashwant M. Deco  
; APPLICANT: Joel Goldstein  
; APPLICANT: Robert Graziano  
; APPLICANT: Chezia Somsandaram  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED  
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/661,052  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/484,172  
; FILING DATE: 07-JUNE-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Arnold, Beth E.  
REGISTRATION NUMBER: 35,430  
REFERENCE/DOCKET NUMBER: MXI-043CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: Internal  
US-08-661-052-12

Query Match 60.2%; Score 112; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21  
DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 5  
US-08-460-502-8  
Sequence 8, Application US/08460502  
Patent No. 5843464  
GENERAL INFORMATION:  
APPLICANT: Bakaletz, Lauren O.  
APPLICANT: Kaumaya, Parvin T.  
TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Calfee, Halter and Griswold  
STREET: 800 Superior Avenue  
CITY: Cleveland  
STATE: Ohio  
COUNTRY: U.S.A.  
ZIP: 44114-2688  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,502  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Goltick, Mary E.  
REGISTRATION NUMBER: 34,829  
REFERENCE/DOCKET NUMBER: 22727/00120  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (216) 622-8458  
TELEFAX: (216) 241-0816  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-502-8

Query Match 60.2%; Score 112; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21  
DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 6  
US-08-724-774B-5  
Sequence 5, Application US/08724774B  
Patent No. 5908778  
GENERAL INFORMATION:  
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,  
APPLICANT: Pierre; Cerrotini, Jean-Charles; Carrel,  
APPLICANT: Stefan; Reed, Daryl  
TITLE OF INVENTION: MAGE-10 ENCODING CDNA, The Tumor  
TITLE OF INVENTION: Rejection Antigen Precursors MAGE-10,  
TITLE OF INVENTION: Antibodies Specific To The Molecule, and  
TITLE OF INVENTION: Uses Thereof  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felle & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,774B  
FILING DATE: 03-October-1996  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, NO. 5908778man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5457  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-724-774B-5

Query Match 60.2%; Score 112; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21  
DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 7  
US-09-089-595-5  
Sequence 5, Application US/09089595  
Patent No. 6153728  
GENERAL INFORMATION:  
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,  
APPLICANT: Pierre; Cerrotini, Jean-Charles; Carrel,  
APPLICANT: Stefan; Reed, Daryl  
TITLE OF INVENTION: MAGE-10 ENCODING CDNA, The Tumor  
TITLE OF INVENTION: Rejection Antigen Precursors MAGE-10,  
TITLE OF INVENTION: Antibodies Specific To The Molecule, and  
TITLE OF INVENTION: Uses Thereof  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felle & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA

ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 Kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/089,595  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,774  
FILING DATE: 03-October-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6153728man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5457  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-089-595-5

Query Match 60.2%; Score 112; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21  
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 8  
US-09-382-855-5  
Sequence 5, Application US/09382855  
Patent No. 6174692  
GENERAL INFORMATION:  
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulle, Pierre;  
TITLE OF INVENTION: Mage-10 ENCODING cDNA, The Tumor Rejection  
TITLE OF INVENTION: Antigen Precursors Mage-10, Antibodies Specific To The Molecule,  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESS: Fulbright & Jaworski LLP  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/382,855  
FILING DATE: 25-August-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/089,595  
FILING DATE: 02-June-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,774  
FILING DATE: 03-October-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Schofield, Mary Anne  
REGISTRATION NUMBER: 36,669  
REFERENCE/DOCKET NUMBER: LUD 5457.2 DIV - JEL/MAS  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 318-3000  
TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-382-855-5

Query Match 60.2%; Score 112; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21  
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 9  
US-09-183-714B-5  
Sequence 5, Application US/09183714B  
Patent No. 6221593  
GENERAL INFORMATION:  
APPLICANT: Boon-Falleur, Thierry  
APPLICANT: Brasseur, Francis  
APPLICANT: Rimoldi, Donata  
APPLICANT: De Plaen, Etienne  
TITLE OF INVENTION: Method for Determining Cancer by Determining Expression  
TITLE OF INVENTION: Of Mage-10  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/09/183,714B  
CURRENT FILING DATE: 1998-10-30  
PRIOR APPLICATION NUMBER: US 08/724,774  
PRIOR FILING DATE: 1996-10-03  
NUMBER OF SEQ ID NOS: 7  
SEQ ID NO 5  
LENGTH: 21  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-183-714B-5

Query Match 60.2%; Score 112; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21  
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 10  
US-09-188-082-12  
Sequence 12, Application US/09188082  
Patent No. 6270765  
GENERAL INFORMATION:  
APPLICANT: Yashwant M. Deo  
APPLICANT: Joel Goldstein  
APPLICANT: Robert Graziano  
APPLICANT: Chezian Somasundaram  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISD  
TITLE OF INVENTION: OF ANTI-PC RECEPTOR ANTIBODIES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESS: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/188,082  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/661,052  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Arnold, Beth E.  
REGISTRATION NUMBER: 35,430  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-09-188-082-12

Query Match 60.2%; Score 112; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSEFWLRVPKVSASHLE 21  
DB 1 FNNFTVSEFWLRVPKVSASHLE 21

RESULT 11  
US-09-171-969-10  
Sequence 10, Application US/09171969  
Patent No. 6284533  
GENERAL INFORMATION:  
APPLICANT: Thomas, Lawrence J.  
TITLE OF INVENTION: PLASMD-BASED VACCINE FOR TREATING ATHEROSCLEROSIS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 75 State Street, Suite 2300  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1807  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/171,969  
FILING DATE: 01 May 1997 (01.05.97)  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/640,713  
FILING DATE: 01 May 1996 (01.05.96)  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/802,967  
FILING DATE: 21 February 1997 (21.02.97)  
ATTORNEY/AGENT INFORMATION:  
NAME: Leon R. Yankwich  
REGISTRATION NUMBER: 30,237  
REFERENCE/DOCKET NUMBER: TCS 414.1 PCT (05872)  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

HYPOTHETICAL:  
ANTI-SENSE:  
FEATURE:  
NAME/KEY:  
LOCATION:  
US-09-171-969-10

Query Match 60.2%; Score 112; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSEFWLRVPKVSASHLE 21  
DB 1 FNNFTVSEFWLRVPKVSASHLE 21

RESULT 12  
US-09-364-088-12  
Sequence 12, Application US/09364088  
Patent No. 6365161  
GENERAL INFORMATION:  
APPLICANT: Yashwant M. Deo, et al.  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED  
OF ANTI-FC RECEPTOR ANTIBODIES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street, 24th Floor  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/364,088  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/188,082  
FILING DATE: 07-JUNE-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/484,172  
FILING DATE: 07-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Remillard, Jane E.  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: MXI-043CP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)742-7414  
TELEFAX: (617)227-7400  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-09-364-088-12

Query Match 60.2%; Score 112; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSEFWLRVPKVSASHLE 21  
DB 1 FNNFTVSEFWLRVPKVSASHLE 21

RESULT 13  
US-09-642-281-5

Sequence 5, Application US/09642281  
Patent No. 6387698  
GENERAL INFORMATION:  
APPLICANT: Rimoldi, Donata; Tongeneel, Victor; Coullie, Pierre;  
Carrotlin, Jean-Charles; Carrel, Stefan; Reed, Daryl  
TITLE OF INVENTION: MAGE-10 ENCODING CDNA, The Tumor Rejection  
Antigen Precursors MAGE-10, Antibodies Specific To The Molecu  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski LLP  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/642,281  
FILING DATE: 18-Aug-2000  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/089,595  
FILING DATE: 02-June-1998  
APPLICATION NUMBER: US 08/724,774  
FILING DATE: 03-October-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Schofield, Mary Anne  
REGISTRATION NUMBER: 36,669  
REFERENCE/DOCKET NUMBER: LUD 5457.2 DIV - JEL/MAS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3000  
TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-642-281-5  
Query Match 60.2%; Score 112; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FNNFTVSFWLRVPKVSASHLE 21  
DB 1 FNNFTVSFWLRVPKVSASHLE 21  
RESULT 14  
US-09-102-716-12  
Sequence 12, Application US/09102716  
Patent No. 6395272  
GENERAL INFORMATION:  
APPLICANT: Yashwant M. Deo  
Joel Goldstein  
Robert Graziano  
Chezian Somasundaram  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISD  
OF ANTI-FC RECEPTOR ANTIBODIES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/102,716  
FILING DATE: 22-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/661,052  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Arnold, Beth E.  
REGISTRATION NUMBER: 35,430  
REFERENCE/DOCKET NUMBER: MXI-043CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-102-716-12  
Query Match 60.2%; Score 112; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FNNFTVSFWLRVPKVSASHLE 21  
DB 1 FNNFTVSFWLRVPKVSASHLE 21  
RESULT 15  
US-08-432-483A-3  
Sequence 3, Application US/08432483A  
Patent No. 6410022  
GENERAL INFORMATION:  
APPLICANT: Rittershaus, Charles W.  
TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER  
TRANSFER PROTEIN (CETP) ACTIVITY  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Wilcoff, Ltd.  
STREET: Ten South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-7407  
COMPUTER READABLE FORM:  
MEDIUM TYPE: diskette  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/432,483A  
FILING DATE: 1-May-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Leon R. Yankwich  
REGISTRATION NUMBER: 30,237  
REFERENCE/DOCKET NUMBER: 95,179(TCS-95179)  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
HYPOTHETICAL:  
ANTI-SENSE:  
FEATURE:  
NAME/KEY: 21-amino acid tetanus toxoid universal  
NAME/KEY: helper T cell epitope.  
LOCATION:  
PUBLICATION INFORMATION:  
AUTHORS: Pantina-Bordignon, P., et al.  
TITLE: Universally immunogenic T cell  
TITLE: epitopes: promiscuous binding to human MHC class II and  
JOURNAL: Promiscuous recognition by T cells  
JOURNAL: European Journal of Immunology  
VOLUME: 19  
ISSUE:  
PAGES: 2237-2242  
DATE: 1989  
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 21  
US-08-432-483A-3

Query Match 60.2%; Score 112; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. NO. 6.8e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKYSASHLE 21  
|||||  
Db 1 FNNFTVSFWLRVPKYSASHLE 21

Search completed: September 24, 2003, 17:00:22  
Job time : 11.1776 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:57:10 : Search time 17.0471 seconds  
(without alignments)  
301.778 Million cell updates/sec

Title: US-09-848-834A-10  
Perfect score: 186  
Sequence: 1 FNNFTVSFWLRPKVKSASHLEGPLHMSYGLRXP 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185	99.5	34	9	US-09-848-834A-10
2	185	99.5	50	9	US-09-848-834A-18
3	112	60.2	21	9	US-09-943-548-3
4	112	60.2	21	9	US-09-848-834A-4
5	112	60.2	21	10	US-09-785-215-6
6	112	60.2	21	12	US-09-405-986-2
7	112	60.2	21	12	US-10-223-809A-6
8	112	60.2	21	12	US-10-261-208-5
9	112	60.2	21	15	US-10-204-362-6
10	112	60.2	21	15	US-10-339-522-3
11	112	60.2	21	15	US-10-223-711-8
12	112	60.2	37	9	US-09-848-834A-14
13	112	60.2	463	11	US-09-816-467-2
14	112	60.2	605	12	US-10-130-973A-11
15	112	60.2	665	12	US-10-130-973A-9

16	112	60.2	882	12	US-10-130-973A-3	Sequence 3, Appl
17	112	60.2	907	12	US-10-130-973A-5	Sequence 5, Appl
18	112	60.2	1052	12	US-10-130-973A-17	Sequence 17, Appl
19	112	60.2	1112	12	US-10-130-973A-16	Sequence 16, Appl
20	112	60.2	1315	12	US-10-241-596-141	Sequence 141, App
21	107	57.5	22	11	US-09-865-294-4	Sequence 4, Appl
22	95.5	51.3	20	14	US-10-044-034-20	Sequence 20, Appl
23	88	47.3	22	11	US-09-747-802-12	Sequence 12, Appl
24	79	42.5	14	10	US-09-909-460-43	Sequence 43, Appl
25	79	42.5	15	10	US-09-909-460-44	Sequence 44, Appl
26	79	42.5	31	9	US-09-848-834A-9	Sequence 9, Appl
27	79	42.5	47	9	US-09-848-834A-17	Sequence 17, Appl
28	73	39.2	28	9	US-09-848-834A-11	Sequence 11, Appl
29	73	39.2	33	9	US-09-848-834A-12	Sequence 12, Appl
30	73	39.2	46	9	US-09-848-834A-19	Sequence 19, Appl
31	73	39.2	51	9	US-09-848-834A-20	Sequence 20, Appl
32	61	32.8	144	8	US-08-981-087A-2	Sequence 2, Appl
33	61	32.8	431	8	US-08-981-087A-1	Sequence 1, Appl
34	61	32.8	432	11	US-09-910-186A-16	Sequence 16, Appl
35	61	32.8	432	11	US-09-910-186A-34	Sequence 34, Appl
36	61	32.8	449	11	US-09-910-186A-18	Sequence 18, Appl
37	61	32.8	645	12	US-10-130-973A-8	Sequence 8, Appl
38	61	32.8	660	12	US-10-130-973A-12	Sequence 12, Appl
39	61	32.8	685	12	US-10-130-973A-7	Sequence 7, Appl
40	61	32.8	862	12	US-10-130-973A-4	Sequence 4, Appl
41	61	32.8	887	12	US-10-130-973A-6	Sequence 6, Appl
42	61	32.8	1032	12	US-10-130-973A-15	Sequence 15, Appl
43	61	32.8	1092	12	US-10-130-973A-14	Sequence 14, Appl
44	59	31.7	49	9	US-09-019-010-4	Sequence 4, Appl
45	59	31.7	49	11	US-09-305-924-11	Sequence 11, Appl

## ALIGNMENTS

```
RESULT 1
US-09-848-834A-10
: Sequence 10, US0020076416A1
: Patent No. US20020076416A1
: GENERAL INFORMATION:
: APPLICANT: Aption Corporation
: TITLE OF INVENTION: Chimeric Peptide Immunogens
: FILE REFERENCE: 1102865-0047
: CURRENT APPLICATION NUMBER: US/09/848,834A
: CURRENT FILING DATE: 2001-05-04
: PRIOR APPLICATION NUMBER: 60/202,328
: PRIOR FILING DATE: 2000-05-05
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 10
: LENGTH: 34
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of
: OTHER INFORMATION: Tetanus toxoid precursor (Tetoxylysin) linked by a spacer to
: OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
: NAME/KEY: MOD_RES
: LOCATION: (1)..(1)
: OTHER INFORMATION: Amidated phenylalanine
: NAME/KEY: PEPTIDE
: LOCATION: (1)..(21)
: OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor
: NAME/KEY: PEPTIDE
: LOCATION: (22)..(25)
: OTHER INFORMATION: (Tetoxylysin)
: NAME/KEY: PEPTIDE
: LOCATION: (26)..(34)
: OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone
: NAME/KEY: MOD_RES
: LOCATION: (34)..(34)
: OTHER INFORMATION: Amidated glycine or glycylamide
```

US-09-848-834A-10

Query Match 99.5%; Score 185; DB 9; Length 34;  
Best Local Similarity 100.0%; Pred. No. 5e-18;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNNFTVSFWLRVPKVSASHLEGPSTLHWSYGLRP 33  
Db 1 FNNFTVSFWLRVPKVSASHLEGPSTLHWSYGLRP 33

RESULT 2

US-09-848-834A-18  
; Sequence 18, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphon Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 18  
; LENGTH: 50  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human GnRH linked by a spacer to amino acid sequence 947-967 of the Tet  
OTHER INFORMATION: anus toxoid precursor (Tentoxylisin) protein linked by a spacer to  
OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH

NAME/KEY: MOD\_RES  
LOCATION: (1)..(1)  
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: MOD\_RES  
LOCATION: (50)..(50)  
OTHER INFORMATION: Amidated glycine or glycylamide

NAME/KEY: PEPTIDE  
LOCATION: (1)..(10)  
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone

NAME/KEY: PEPTIDE  
LOCATION: (11)..(16)  
OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE  
LOCATION: (17)..(37)  
OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tentoxylisin)

NAME/KEY: PEPTIDE  
LOCATION: (38)..(41)  
OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE  
LOCATION: (42)..(50)  
OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone

US-09-848-834A-18

Query Match 99.5%; Score 185; DB 9; Length 50;  
Best Local Similarity 100.0%; Pred. No. 7.6e-18;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNNFTVSFWLRVPKVSASHLEGPSTLHWSYGLRP 33  
Db 17 FNNFTVSFWLRVPKVSASHLEGPSTLHWSYGLRP 49

RESULT 3

US-09-943-548-3  
; Sequence 3, Application US/09943548  
; Patent No. US20020042364A1  
; GENERAL INFORMATION:  
; APPLICANT: Rittershaus, Charles W.

APPLICANT: Thomas, Lawrence J.  
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CEMP) ACTIVITY  
; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2  
; CURRENT APPLICATION NUMBER: US/09/943,548  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 08/432,483  
; PRIOR FILING DATE: 1995-05-01  
; PRIOR APPLICATION NUMBER: PCT/US96/06147  
; PRIOR FILING DATE: 1996-05-01  
; PRIOR APPLICATION NUMBER: 08/945,289  
; PRIOR FILING DATE: 1997-10-17  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

OTHER INFORMATION: helper T cell epitope of tetanus toxin

Query Match 60.2%; Score 112; DB 9; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.1e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNNFTVSFWLRVPKVSASHLE 21  
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 4

US-09-848-834A-4  
; Sequence 4, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphon Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Tetanus bacillus

NAME/KEY: PEPTIDE  
LOCATION: (1)..(21)  
OTHER INFORMATION: Amino acid sequence 947-967 of Tetanus

OTHER INFORMATION: Toxoid Precursor (Tentoxylisin)

US-09-848-834A-4

Query Match 60.2%; Score 112; DB 9; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.1e-08;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNNFTVSFWLRVPKVSASHLE 21  
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 5

US-09-785-215-6  
; Sequence 6, Application US/09785215  
; Publication No. US20020187157A1  
; GENERAL INFORMATION:  
; APPLICANT: JENSEN, Martin Roland et al.  
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID  
; FILE REFERENCE: 3631-0107P  
; CURRENT APPLICATION NUMBER: US/09/785,215

```
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 21
; TYPE: PRF
; ORGANISM: Clostridium tetani
US-09-785-215-6

Query Match
Best Local Similarity 100.0%; Score 112; DB 10; Length 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNNFTVSEFWLRVPKVSASHLE 21
DB 1 FNNFTVSEFWLRVPKVSASHLE 21

RESULT 6
US-09-405-986-2
; Sequence 2, Application US/09405986
; Publication No. US20030157115A1
; GENERAL INFORMATION:
; APPLICANT: BAY, Sylvie
; APPLICANT: CANTACUZE, Daniele
; APPLICANT: LECIERC, Claude
; APPLICANT: LO-MAN, Richard
; TITLE OF INVENTION: MULTIPLE ANTIGEN GLYCOPOLYMER CARBOHYDRATE, VACCINE
; FILE REFERENCE: 1341 US 3565
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: 60/041,726
; EARLIER FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 21
; TYPE: PRF
; ORGANISM: Clostridium tetani
US-09-405-986-2

Query Match
Best Local Similarity 100.0%; Score 112; DB 12; Length 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNNFTVSEFWLRVPKVSASHLE 21
DB 1 FNNFTVSEFWLRVPKVSASHLE 21

RESULT 7
US-10-223-809A-6
; Sequence 6, Application US/10223809A
; Publication No. US20030157117A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; APPLICANT: Rasmussen, Peter Birk et al.
; TITLE OF INVENTION: No. US20030157117A1 Method for Down-Regulation of Amyloid
; FILE REFERENCE: 674542-2008
; CURRENT FILING DATE: 2002-08-20
; CURRENT APPLICATION NUMBER: US/10/223,809A
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 60/337,543
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 60/373,027
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: DE 2001 01231
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: DE 2002 0058
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
```

```
; LENGTH: 21
; TYPE: PRF
; ORGANISM: Clostridium tetani
US-10-223-809A-6

Query Match
Best Local Similarity 100.0%; Score 112; DB 12; Length 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNNFTVSEFWLRVPKVSASHLE 21
DB 1 FNNFTVSEFWLRVPKVSASHLE 21

RESULT 8
US-10-261-208-5
; Sequence 5, Application US/10261208
; Publication No. US20030158388A1
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulle,
; Pierre; Cerrotini, Jean-Charles; Carrel,
; Stefan; Reed, Daryl
; TITLE OF INVENTION: MAGE-10 ENCODING CDNA, The Tumor
; Rejection Antigen Precursors MAGE-10,
; Antibodies Specific To The Molecule, and
; Uses Thereof
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/261,208
; FILING DATE: 30-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/589,717
; FILING DATE: 08-Jun-2000
; APPLICATION NUMBER: 09/089,595
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20030158388A1man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-261-208-5

Query Match
Best Local Similarity 100.0%; Score 112; DB 12; Length 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FNNFTVSEFWLRVPKVSASHLE 21
DB 1 FNNFTVSEFWLRVPKVSASHLE 21

RESULT 9
```

```
US-10-204-362-6
; Sequence 6, Application US/10204362
; Publication No. US20030086938A1
; GENERAL INFORMATION:
; APPLICANT: M&E Biotech A/S
; TITLE OF INVENTION: No. US20030086938A1 Method For Down-Regulation Of Amyloid
; FILE REFERENCE: 3631-0120P
; CURRENT APPLICATION NUMBER: US/10/204,362
; CURRENT FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-204-362-6

Query Match          60.2%; Score 112; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 2,1e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
   |||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 10
US-10-339-522-3
; Sequence 3, Application US/10339522
; Publication No. US20030106559A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CEMP) ACTIVITY
; FILE REFERENCE: TCS-411.1P US-3
; CURRENT APPLICATION NUMBER: US/10/339,522
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 09/943,334
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/943,548
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: helper T cell epitope of tetanus toxin
US-10-339-522-3

Query Match          60.2%; Score 112; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 2,1e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
   |||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 11
US-10-223-711-8
; Sequence 8, Application US/10223711
; Publication No. US20030113344A1
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Pravin T.P.
US-10-223-711-8

; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; FILE REFERENCE: 18525/04058
; CURRENT APPLICATION NUMBER: US/10/223,711
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 09/148,711
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 08/460,502
; PRIOR FILING DATE: 1995-06-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-10-223-711-8

Query Match          60.2%; Score 112; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 2,1e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
   |||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 12
US-09-848-834A-14
; Sequence 14, Application US/0984834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
; OTHER INFORMATION: GnRH hormone linked by a spacer to amino acid sequence 947-967
; OTHER INFORMATION: the tetanus toxoid precursor (Tentoxylisin)
; NAME/KEY: MOD-RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(37)
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor
; OTHER INFORMATION: (Tentoxylisin)
US-09-848-834A-14

Query Match          60.2%; Score 112; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 3,8e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
   |||||
Db 17 FNNFTVSFWLRVPKVSASHLE 37

RESULT 13
US-09-816-467-2
```

```
; Sequence 2, Application US/09816467
; Publication No. US20030004121A1
; GENERAL INFORMATION:
; APPLICANT: COEN, LAURENT
; APPLICANT: PINZOLAS, ROSARIO OSTA
; APPLICANT: BRULEY, PHILIPPE
; TITLE OF INVENTION: HYBRID PROTEINS THAT MIGRATE RETROGRADELY AND
; TITLE OF INVENTION: TRANSSYNAPTICALLY INTO THE CNS
; FILE REFERENCE: 03495.0174-01000
; CURRENT APPLICATION NUMBER: US/09/816.467
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/055,615
; PRIOR FILING DATE: 1997-08-14
; PRIOR APPLICATION NUMBER: 60/065,236
; PRIOR FILING DATE: 1997-11-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-816-467-2

Query Match          60.2%; Score 112; DB 11; Length 463;
Best Local Similarity 100.0%; Pred. No. 5.5e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
DB 95 FNNFTVSFWLRVPKVSASHLE 115

RESULT 14
US-10-130-973A-11
; Sequence 11, Application US/10130973A
; Publication No. US20030147895A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford
; APPLICANT: Sutton, John
; APPLICANT: Silman, Nigel
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cells
; FILE REFERENCE: 1581.0920000
; CURRENT APPLICATION NUMBER: US/10/130.973A
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: PCT/GB00/04644
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: GB 9928530.6
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: GB 008658.7
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-130-973A-11

Query Match          60.2%; Score 112; DB 12; Length 605;
Best Local Similarity 100.0%; Pred. No. 7.3e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
DB 235 FNNFTVSFWLRVPKVSASHLE 255

RESULT 15
US-10-130-973A-9
; Sequence 9, Application US/10130973A
; Publication No. US20030147895A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford
; APPLICANT: Sutton, John
; APPLICANT: Silman, Nigel
; TITLE OF INVENTION: Constructs for Delivery of Therapeutic Agents to Neuronal Cell
; FILE REFERENCE: 1581.0920000
; CURRENT APPLICATION NUMBER: US/10/130.973A
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: PCT/GB00/04644
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: GB 9928530.6
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: GB 008658.7
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-130-973A-9

Query Match          60.2%; Score 112; DB 12; Length 665;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
DB 295 FNNFTVSFWLRVPKVSASHLE 315
```

Search completed: September 24, 2003, 17:33:07  
Job time : 18.0471 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:30:06 ; Search time 11.9612 Seconds  
(without alignments)  
273.361 Million cell updates/sec

Title: US-09-848-834A-10

Perfect score: 186

Sequence: 1 FNNFTVSFWLRAVPKVSASHLGGPSLHMSYGLRFX 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	60.2	1315	1 BTCLTN	tentoxilysin (EC 3
2	62	33.3	1268	2 S33411	botulinum neurotox
3	61.5	33.1	92	1 RHRTG	gonadoliberin prec
4	61	32.8	366	2 S48110	neurotoxin type F
5	61	32.8	369	2 S48109	neurotoxin type F
6	61	32.8	1274	2 I40813	neurotoxin type F
7	61	32.8	1297	2 S39791	neurotoxin - Clost
8	59	31.7	1296	1 BTCLAB	botoxilysin (EC 3
9	58	31.2	1291	1 A48940	botoxilysin (EC 3
10	58	31.2	1291	1 I40631	non-protolytic ho
11	57.5	30.9	90	1 RHMSG	gonadoliberin prec
12	57	30.6	502	2 T36589	probable transmem
13	56.5	30.4	367	2 S46106	neurotoxin type E
14	56.5	30.4	1251	2 JH0256	botulinum neurotox
15	56.5	30.4	1252	2 S21178	botulinum neurotox
16	56	30.1	812	2 T01618	hypothetical prote
17	56	30.1	1296	2 I40645	botulinum neurotox
18	54.5	29.3	92	1 RHMSG	gonadoliberin prec
19	54.5	29.3	1285	2 S70582	botulinum neurotox
20	54.5	29.3	1291	2 S46431	botulinum neurotox
21	54.5	29.3	1291	2 A49777	botulinum neurotox
22	54	28.0	67	2 I78541	gonadoliberin prec
23	53.5	28.8	469	2 B37837	probable alpha-am
24	53.5	28.8	3122	2 T17202	DNA-directed DNA p
25	53	28.5	519	2 S78196	probable maturase
26	52	28.0	10	1 RHSGG	gonadoliberin - pl
27	52	28.0	10	1 RHSGG	gonadoliberin - sh
28	52	28.0	89	2 I51423	gonadoliberin prec
29	52	28.0	449	2 S23158	nucleocapsid prote

30	52	28.0	464	1 MNVUMC
31	52	28.0	467	1 MNVUM1
32	51.5	27.7	91	2 UC7393
33	51.5	27.7	537	2 S78195
34	51.5	27.7	573	2 S78197
35	51	27.4	292	2 S22441
36	51	27.4	496	2 T38197
37	51	27.4	659	2 F70453
38	51	27.4	944	2 T18627
39	51	27.4	1196	2 UQ1467
40	51	27.4	1196	2 S46430
41	50.5	27.2	438	2 T52082
42	50.5	27.2	1450	2 T30273
43	50	26.9	432	2 T04726
44	50	26.9	591	2 S04401
45	50	26.9	836	2 JE0248

## ALIGNMENTS

RESULT 1  
BTCLTN  
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani  
N:Alternate names: tetanus neurotoxin  
C:Species: Clostridium tetani  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 03-Jun-2002  
C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364  
R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, EMO J. 5, 2495-2502, 1986  
A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b  
A:Reference number: A25689; MUID:87053814; PMID:3536478  
A:Accession: A25689  
A:Residues: 1-1315 <EIS>  
A:Molecule type: DNA  
R:Fairweather, N.F.; Lyness, V.A.  
Nucleic Acids Res. 14, 7809-7812, 1986  
A:Title: The complete nucleotide sequence of tetanus toxin.  
A:Reference number: A25757; MUID:87040747; PMID:3774547  
A:Accession: A25757  
A:Molecule type: DNA  
A:Residues: 1-1315 <FAI>  
A:Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:940774  
A:Experimental source: Strain CN3911  
R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.  
J. Bacteriol. 165, 21-27, 1986  
A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C 1  
A:Reference number: A25194; MUID:86085672; PMID:3510187  
A:Accession: A25194  
A:Molecule type: DNA  
A:Residues: 743-1315 <PA2>  
A:Cross-references: GB:M12739; NID:944920; PIDN:AAA23282.1; PID:944921  
A:Accession: B25194  
A:Molecule type: protein  
A:Residues: 865-894 <PA3>  
R:Matsumoto, M.; Lei, D.U.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.  
Infect. Immun. 57, 3588-3593, 1989  
A:Title: Isolation, purification, and characterization of fragment B, the NH-2-termin  
A:Reference number: A60759; MUID:90035436; PMID:2478476  
A:Accession: A60759  
A:Molecule type: protein  
A:Residues: 461-475 <MAT>  
R:Demetz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.  
J. Immunol. 142, 394-402, 1989  
A:Title: Delination of several DR-restricted tetanus toxin T cell epitopes.  
A:Reference number: J50098; MUID:89093918; PMID:2463305  
A:Contents: annotation; epitope region  
R:Schlavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta, Nature 359, 832-835, 1992  
A:Title: Tetanus and Botulinum-B neurotoxins block neurotransmitter release by proteo  
A:Reference number: S27125; MUID:93063293; PMID:1331807  
A:Contents: annotation

E:Ride, Filippos V., Vangelista, L., Schlavo, G., Tonello, F., Montecucco, C.  
 Eur. J. Biochem. 229, 61-69, 1995  
 A:Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin  
 A:Reference number: S69348; MUID:95262688; PMID:7744050  
 A:Accession: S69348  
 A:Molecule type: protein  
 A:Residues: 2-31 <DEF>  
 C:Comment: The source of this protein was an extrachromosomal plasmid.  
 C:Comment: The precursor is cleaved by endogenous proteinase activity to form light (fragment A) and heavy (fragment B) chains.  
 C:Comment: The amino end of the heavy chain (fragment B) forms ion channels in a lipid bilayer. Fragment C binds to ganglioside GM1 and is internalized by endocytosis.  
 C:Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized by endocytosis, and inhibits neurotransmitter release by proteolytic cleavage of synaptic function.  
 A:Description: blocks neuroexocytosis via hydrolysis of a Glu-Phe peptide bond in synaptobrevin.  
 C:Superfamily: tetanus toxin  
 C:Keywords: hydrolyase; metalloproteinase; neurotoxin; transmembrane protein; zinc  
 F:2-457/Product: tenotoxylsln light chain (fragment A) #status predicted <TTH>  
 F:461-1315/Product: tenotoxylsln heavy chain (fragment B.C) #status experimental <TTH>  
 F:461-864/Domian: channel forming (fragment B) #status predicted <TTH>  
 F:865-1315/Domian: ganglioside binding (fragment C) #status predicted <TTH>  
 F:233,237/Binding site: zinc (His) #status predicted  
 F:234/Active site: Glu #status predicted

```

Query Match          60.2%   Score 112   DB 1; Length 1315;
Best Local Similarity 100.0%; Prod. No. 1_9e-07;
Matches      21; Conservative    0; Mismatches     0; Indels     0; Gaps     0;
OY           1 FNNFTVSFWLVRPKVSASHLE 21
              |||
db            947 FNNFTVSFWLVRPKVSASHLE 967

```

```

RESULT 2
S33411
botulinum neurotoxin type F - Clostridium barati
C:Species: Clostridium barati
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C:Accession: S33411; S31860
R:Thompson, D.E.; Huxson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T
PMNS Microbiol. Lett. 108, 175-182, 1993
A:Title: Nucleotide sequence of the gene coding for Clostridium barati type F neurotoxin
A:Reference number: S33411; MUID:93252228; PMID:8486245
A:Accession: S33411
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1268 <THO>
A:Cross-references: EMBL:X68262; NID:949138; PIDN:CA048329.1; PID:949139
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

```

Query Match	33.3%	Score 62	DB 2	Length 1268
Best Local Similarity	64.3%	Pred. No. 2.2		
Matches	9	Conservative	4	Mismatches
			1	Indels
			0	Gaps
0y	1	FNNFTVSFMLEVPK	14	
db	922	YQNFVSFMRVRIK	935	

RESULT 3  
RHRTG  
gonadoliberin precursor - rat  
N:Alternate names: gonadotropin-associated protein (GAP); gonadotropin releasing hormone  
N:Contents: gonadoliberin; prolactin release-inhibiting factor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1988 #sequence, revision 31-Mar-1988 #text, change 18-Jun-1999  
C:Accession: A40147; B26173; A48410  
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.  
Mol. Endocrinol. 3, 1257-1262, 1989  
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex  
A:Reference number: A40147; MUID:89384661; PMID:2476669  
I:Accession: A40147

A:Molecule type: DNA  
A:Residues: 1-92 <BON>  
A:Cross-references: GB:M1670; NID:g204447; PID:AAA41264.1; PID:g204448  
R:Adelman, J.P.; Mason, A.J.; Haylick, J.S.; Seeburg, P.H.  
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986  
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and growth hormone releasing hormone  
A:Reference number: A94090; MUID:86094338; PMID:2867548  
A:Accession: B26173  
A:Molecule type: mRNA  
A:Residues: 1-92 <ADE>  
A:Cross-references: GB:M12579; NID:g204445; PID:AAA41263.1; PID:g204446  
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Bialock, P.H.  
Cell. Mol. Neurobiol. 12, 447-454, 1992  
A:Title: Thyrocytes express a mRNA that is identical to hypothalamic luteinizing hormone-releasing hormone  
A:Reference number: A48410; MUID:93105480; PMID:1468115  
A:Accession: A48410  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-92 <MAI>  
A:Cross-references: GB:S50870; NID:g262059; PID:AAAB24572.1; PID:g262060  
A:Experimental source: thymus  
A:Note: sequence extracted from NCBI backbone (NCB1N:121082, NCB1P:121083)  
C:Genetics:  
A:Introns: 47/3; 79/3  
C:Function:  
A:Description: stimulates pituitary secretion of luteotropin and follitropin  
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity  
C:Superfamily: gonadoliberin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyrogutamic acid  
F:1-33/Domian: signal sequence #status predicted <SIG>  
F:24-92/Product: progondoliberin #status predicted <GN>  
F:24-33/Product: gonadoliberin #status predicted <GN>  
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>  
F:24/Modified site: pyroridone carboxylic acid (Gln) (in mature form) #status predicted <MOD>  
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form) from following

Query Match	33.1%	Score	61.5	DB	1	length	92
Best Local Similarity	51.7%	Pred	No. 0.14				
Matches	15	Conservative	3	Mismatches	4	Indels	7
				Gaps	2		
QY	12	VPKVAS-----HLEG-PSLHWSTGLRP	33				
		: :					
Db	4	IPKIMAAVLLTVLCGGSSQHWSTGLRP	32				

RESULT 4  
S48110  
neurotoxin type F - Clostridium botulinum (fragment)  
C:Species: Clostridium botulinum  
C:Date: 14-Jul-1995 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999  
C:Accession: S48110  
R:Campbell, K.D.; Collins, M.D.; East, A.K.  
J Clin Microbiol. 31, 2255-2262, 1993  
A:Title: Gene probes for identification of the botulinum neurotoxin gene a  
A:Reference number: S48103; MUID:94013372; PMID:8408342  
A:Accession: S48110  
A:Status: Preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-366 <CAM>  
A:Cross-references: EMBL:X70821; NID:g407792; PIDN:CAA50152.1; PID:g4077933  
C:Superfamily: tectanus toxin  
C:Keywords: neurotoxin

Query Match	32.8%	Score 61	DB 2	Length 366
Best Local Similarity	57.1%	Pred No	0.76	
Matches	8	Conservative	5	Mismatches
			1	Indels
				Gaps
				0.
QY	1	FNFTVSWLRVPK	14	
		Y:::    :		
Db	297	YONFSISFWVRPK	310	

## RESULT 5

S48109  
neurotoxin type F - Clostridium botulinum (fragment)  
C:Species: Clostridium botulinum  
C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 16-Jul-1999  
C:Accession: S48109  
R:Campbell, K.D.; Collins, M.D.; East, A.K.  
J. Clin. Microbiol. 31, 2255-2262, 1993  
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific id  
A:Reference number: S48103; M0ID:94013372; PMID:8408542  
A:Accession: S48109  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-369 <CAM>  
A:Cross-references: EMBL:X70820; NID:g407790; PIDN:CAA50151.1; PID:g407791  
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1993  
C:Superfamily: tetanus toxin

Query Match 32.8%; Score 61; DB 2; Length 369;  
Best Local Similarity 57.1%; Pred. No. 0.77;  
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 FNNFTVSEWLRVPK 14  
: |||:||||:|  
Db 297 YQNFSTFWVRIPK 310

RESULT 6  
I40813  
neurotoxin type F - Clostridium botulinum  
C:Species: Clostridium botulinum  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 16-Jul-1999  
C:Accession: I40813; S48108  
R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, D.H.  
FEBS Microbiol. Lett. 96, 225-230, 1992  
A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.  
A:Reference number: I40644  
A:Accession: I40813  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1274 <RES>  
A:Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867  
R:Campbell, K.D.; Collins, M.D.; East, A.K.  
J. Clin. Microbiol. 31, 2255-2262, 1993  
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific id  
A:Reference number: S48103; M0ID:94013372; PMID:8408542  
A:Accession: S48108  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 634-1002 <CAM>  
A:Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789  
C:Superfamily: tetanus toxin  
C:Keywords: neurotoxin

Query Match 32.8%; Score 61; DB 2; Length 1274;  
Best Local Similarity 57.1%; Pred. No. 3;  
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 FNNFTVSEWLRVPK 14  
: |||:||||:|  
Db 930 YQNFSTFWVRIPK 943

RESULT 7  
S39791  
neurotoxin - Clostridium botulinum  
C:Species: Clostridium botulinum  
C:Date: 07-Oct-1994 #sequence\_revision 01-Dec-1995 #text\_change 16-Jul-1999  
C:Accession: S39791  
R:Campbell, K.; Collins, M.D.; East, A.K.  
Biochim. Biophys. Acta 1216, 487-491, 1993  
A:Title: Nucleotide sequence of the gene coding for Clostridium botulinum (Clostridium a  
A:Reference number: S39791; M0ID:94092745; PMID:8268233  
A:Accession: S39791

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1297 <CAM>  
A:Cross-references: EMBL:X74162; NID:g441275; PIDN:CAA52275.1; PID:g441276  
C:Superfamily: tetanus toxin  
C:Keywords: neurotoxin

Query Match 32.8%; Score 61; DB 2; Length 1297;  
Best Local Similarity 38.1%; Pred. No. 3.1;  
Matches 8; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

OY 1 FNNFTVSEWLRVPKVSASHLE 21  
: |||:||||:|  
Db 930 FDNFTVSEWLRVPKYNNDIQ 950

RESULT 8  
BTCLAB  
bontoxillysin (EC 3.4.24.69) A precursor - Clostridium botulinum  
N:Alternate names: botulinum neurotoxin type A  
C:Species: Clostridium botulinum  
C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Jun-1999  
C:Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000  
R:Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.  
J. Biol. Chem. 265, 9153-9158, 1990  
A:Title: The complete sequence of botulinum neurotoxin type A and comparison with ocl  
A:Reference number: A35294; M0ID:90264400; PMID:2160960  
A:Accession: A35294  
A:Molecule type: DNA  
A:Residues: 1-1296 <BIN>  
A:Cross-references: GB:M30196; NID:g144864; PIDN:AAA23262.1; PID:g144865  
A:Experimental source: strain 62A, subtype A  
R:Thompson, D.E.; Brehm, J.K.; Outram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson,  
Eur. J. Biochem. 189, 73-81, 1990  
A:Title: The complete amino acid sequence of the Clostridium botulinum type A neurotc  
A:Reference number: S09492; M0ID:90235864; PMID:2185020  
A:Accession: S09492  
A:Molecule type: DNA  
A:Residues: 1'Q', 3'-26, 'V', 28-1296 <THO>  
A:Cross-references: EMBL:X52066; NID:g40381; PIDN:CAA36289.1; PID:g40382  
A:Experimental source: NCTC 2916  
R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.  
FEBS Lett. 376, 41-44, 1995  
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin compone  
A:Reference number: S67988; M0ID:96096783; PMID:8521962  
A:Accession: S68220  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-12 <FUJ>  
A:Cross-references: EMBL:D67030; DBJ:D50421; NID:g2160224  
R:Beatty, M.J.; Somers, E.; Dasgupta, B.R.  
Biochem. Biophys. Res. Commun. 167, 1388-1395, 1999  
A:Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-t  
A:Reference number: A33401; M0ID:89350959; PMID:2669749  
A:Accession: A33401  
A:Molecule type: DNA  
A:Residues: 1-35 <BET>  
A:Cross-references: GB:M27892; NID:g144880; PIDN:AAA23265.1; PID:g551776  
R:Gimenez, J.A.; Dasgupta, B.R.  
J. Protein Chem. 12, 351-363, 1993  
A:Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42,  
A:Reference number: A53884; M0ID:94000342; PMID:8397793  
A:Accession: A53884  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 867-880, 1148-1217, 'Y', 1219 <GIM>  
A:Experimental source: strain Hall  
A:Note: sequence extracted from NCBI backbone (NCBIP.139159); sequence modified after  
R:Dasgupta, B.R.; Dekleva, M.L.  
Biochimie 72, 661-664, 1990  
A:Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and a  
A:Reference number: A60025; M0ID:91120847; PMID:2126206  
A:Accession: A60025

A:Molecule type: protein  
 A:Residues: 2-6;445-453, 'X', 455-457 <DAS1>  
 R:Dasgupta, B.R.; Foley, J.; Niece, R.  
 Biochemistry 26, 4162, 1987  
 A:Title: Partial sequence of the light chain of botulinum neurotoxin type A.  
 A:Reference number: A27000  
 A:Accession: A27000  
 A:Molecule type: protein  
 A:Residues: 2-47 <DAS2>  
 R:Blitz, T.; Blaszi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;  
 J. Biol. Chem. 269, 1617-1620, 1994  
 A:Title: Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.  
 A:Reference number: A49708; MUID:94124495; PMID:8294407  
 A:Contents: annotation  
 C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synapses  
 C:Genetics:  
 A:Gene: atx; botA  
 C:Function:  
 A:Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associate  
 C:Superfamily: tetanus toxin  
 C:Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane prot  
 F:2-444/Product: botoxylisin A light chain #status experimental <LGHT>  
 F:445-1296/Product: botoxylisin A heavy chain #status experimental <HVT>  
 F:223,227/Binding site: zinc (His) #status predicted  
 F:224/Active site: Glu #status predicted

Query Match 31.7%; Score 59; DB 1; Length 1296;  
 Best Local Similarity 57.1%; Pred. NO. 5.9;  
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 FNNFTVSFWLRVPK 14  
 : ||: ||||: ||||  
 Db 938 YENFTSFWRIRPK 951

## RESULT 9

A48940  
 Botoxylisin (EC 3.4.24.69) B precursor - Clostridium botulinum  
 N:Alternate names: botulinum neurotoxin type B (BoNT/B)  
 C:Species: Clostridium botulinum  
 C:Date: 19-Dec-1993 #sequence revision 18-Nov-1994 #text change 18-Jun-1999  
 R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Manton, N.P.  
 Appl. Environ. Microbiol. 58, 2345-2354, 1992  
 A:Title: Molecular cloning of the Clostridium botulinum structural gene encoding the ty  
 A:Reference number: A48940; MUID:92384550; PMID:1514783  
 A:Accession: A48940  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1291 <WHE>  
 A:Cross-references: GB:M61186; NID:G144734; PIDN:AAA23211.1; PID:G144735  
 A:Experimental source: type B, Danish  
 A:Note: sequence extracted from NCBI backbone (NCBI:112080, NCBIP:112081); this public  
 R:Campbell, K.D.; Collins, M.D.; East, A.K.  
 J. Clin. Microbiol. 31, 2255-2262, 1993  
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id  
 A:Reference number: S48103; MUID:94013372; PMID:8408542  
 A:Accession: S48103  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 634-994 <CAM>  
 A:Cross-references: EMBL:X70817; NID:G407782; PIDN:CAA50148.1; PID:G407783  
 A:Experimental source: proteolytic type B, strain NCTC 7273  
 R:Zabzo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.  
 submitted to the EMBL Data Library, April 1992  
 A:Description: Partial amino acid sequence of botulinum neurotoxin type B and comparisid  
 A:Reference number: S21575  
 A:Accession: S21575  
 A:Molecule type: DNA  
 A:Residues: 36-217, 'G', 219-224, 'S', 226-246 <SZA>  
 A:Cross-references: EMBL:Z11934; NID:G40383; PIDN:CAA77991.1; PID:G40384  
 R:Kurazono, H.; Mochida, S.; Blitz, T.; Elsel, U.; Quanz, M.; Gredenstern, O.; Wernars, K  
 J. Biol. Chem. 267, 14721-14729, 1992

A:Title: Minimal essential domains specifying toxicity of the light chains of tetanus  
 A:Reference number: A42871; MUID:92340509; PMID:1634516  
 A:Accession: A42871  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-313, 'S', 315-451 <KUR>  
 A:Experimental source: strain OKra  
 A:Note: sequence extracted from NCBI backbone (NCBIP:109365)  
 R:Dasgupta, B.R.; Datta, A.  
 Biochimie 70, 811-817, 1988  
 A:Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity w  
 A:Reference number: S07155; MUID:89000887; PMID:3139097  
 A:Accession: S07155  
 A:Molecule type: protein  
 A:Residues: 2-29, 'M', 31-45 <DAS>  
 A:Accession: S08562  
 A:Molecule type: protein  
 A:Residues: 442-463, 'R', 465-467 <DA2>  
 R:Schmidt, J.J.; Sathymoorthy, V.; Dasgupta, B.R.  
 Arch. Biochem. Biophys. 238, 544-548, 1985  
 A:Title: Partial amino acid sequences of botulinum neurotoxins types B and E.  
 A:Reference number: S07128; MUID:85197963; PMID:3888113  
 A:Accession: S07128  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 2-16 <SCH1>  
 A:Accession: S08573  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 2-17 <SCH2>  
 A:Accession: S08574  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 442-459 <SCH3>  
 R:Schavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta,  
 Nature 359, 832-835, 1992  
 A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo  
 A:Reference number: S27125; MUID:93063293; PMID:1331807  
 A:Contents: annotation  
 C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy  
 C:Genetics:  
 A:Gene: Dnct/B  
 C:Function:  
 A:Description: catalyzes hydrolysis of a Glu-Phe peptide bond in synaptobrevin 2  
 C:Superfamily: tetanus toxin  
 C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc  
 F:2-441/Product: botoxylisin B light chain #status experimental <LGHT>  
 F:442-1291/Product: botoxylisin B heavy chain #status experimental <HVT>  
 F:230,234/Binding site: zinc (His) #status predicted  
 F:231/Active site: Glu #status predicted

Query Match 31.2%; Score 58; DB 1; Length 1291;  
 Best Local Similarity 64.3%; Pred. NO. 8.1;  
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 FNNFTVSFWLRVPK 14  
 : ||: ||||: ||||  
 Db 923 FLDFVSFWIRPK 936

## RESULT 10

I40631  
 non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum  
 C:Species: Clostridium botulinum  
 C:Date: 12-Aug-1996 #sequence revision 12-Aug-1996 #text change 16-Jul-1999  
 R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.  
 Curr. Microbiol. 28, 101-110, 1994  
 A:Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botu  
 A:Reference number: I40631; MUID:94122659; PMID:7764370  
 A:Accession: I40631  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA



A:Accession: JH0256  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-27, 'E', 29-1251 <POU>  
 A:Cross-references: EMBL:X62088; NID:940379  
 A:Experimental source: strains ATCC 43181 and ATCC 43755  
 R:Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa, N.  
 J. Gen. Microbiol. 137, 519-525, 1991  
 A:Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E toxin  
 A:Reference number: S16145; MUID:91237316; PMID:2033376  
 A:Accession: S16145  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-229, 'M', 231-252 <PUJ>  
 A:Cross-references: EMBL:X53180; NID:940407; PIDN:CAA37321.1; PID:940408  
 A:Experimental source: strain B16340  
 C:Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release  
 C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the lig  
 C:Superfamily: tetanus toxin  
 C:Keywords: neurotoxin  
 F:2-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>  
 F:423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>  
 F:412-426/Disulfide bonds: #status predicted

Query Match 30.4%; Score 56.5; DB 2; Length 1251;  
 Best Local Similarity 22.4%; Pred. No. 13;  
 Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;

OY 1 FNNFTVSFWLRP-----KVSASHLE-----GP 23  
 :||:||||:|  
 Db 912 YKNSISFWVRIPYDKIKVNNXYTIINCMRDNSGKVSILNHEIITWLDNSGINQ 971

OY 24 SLHWSYG 30  
 |::||  
 Db 972 KLAFTNYG 978

## RESULT 15

S21178  
 botulinum neurotoxin type E precursor - Clostridium botulinum

C:Species: Clostridium botulinum

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 15-Oct-1999

C:Accession: S21178; S48107; JH0257; B35294; A60027; S18111

R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.

Eur. J. Biochem. 204, 657-667, 1992

A:Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin

A:Reference number: S21178; MUID:92174922; PMID:1541280

A:Accession: S21178

A:Molecule type: DNA

A:Residues: 1-1252 <WHE>

A:Cross-references: EMBL:X62683; NID:940397; PIDN:CAA44558.1; PID:940398

R:Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id

A:Reference number: S48103; MUID:9401372; PMID:8408542

A:Accession: S48107

A:Molecule type: DNA

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Residues: 616-982 <CAM>

A:Cross-references: EMBL:X70815; NID:9407786; PIDN:CAA50146.1; PID:9407787

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.

Biochem. Biophys. Res. Commun. 183, 107-113, 1992

A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type

A:Reference number: JH0256; MUID:92181428; PMID:1543481

A:Accession: JH0257

A:Molecule type: DNA

A:Status: nucleic acid sequence not shown

A:Residues: 1-176, 'R', 178-197, 'C', 199-339, 'R', 341-772, 'I', 774-962, 'FE', 965-966, 'R', 968-1

A:Cross-references: EMBL:X63089; NID:940393; PIDN:CAA43999.1; PID:940394

A:Experimental source: strain Beluga

R:Binz, T.; Kuzazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.

A:Title: The complete sequence of botulinum neurotoxin type A and comparison with oth  
 A:Reference number: A35294; MUID:90264400; PMID:2160960  
 A:Accession: B35294  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-176, 'R', 178-252 <BIN>  
 A:Experimental source: strain Beluga  
 R:Gimenez, J.A.; Dasgupta, B.R.  
 Biochimie 72, 213-217, 1990  
 A:Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the  
 A:Reference number: A60027; MUID:90344918; PMID:2116911  
 A:Accession: A60027  
 A:Molecule type: protein  
 A:Residues: 420-427 <GIM>  
 A:Experimental source: strain Beluga  
 A:Note: this fragment was generated by proteolysis with Lys-C rather than with trypsin  
 C:Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit  
 C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the  
 C:Superfamily: tetanus toxin  
 C:Keywords: neurotoxin  
 F:1-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>  
 F:423-1252/Product: botulinum neurotoxin type E heavy chain #status predicted <HCH>  
 F:412-426/Disulfide bonds: #status predicted

Query Match 30.4%; Score 56.5; DB 2; Length 1252;  
 Best Local Similarity 22.4%; Pred. No. 13;  
 Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;

OY 1 FNNFTVSFWLRP-----KVSASHLE-----GP 23  
 :||:||||:|  
 Db 912 YKNSISFWVRIPYDKIKVNNXYTIINCMRDNSGKVSILNHEIITWLDNSGINQ 971

OY 24 SLHWSYG 30  
 |::||  
 Db 972 KLAFTNYG 978

Search completed: September 24, 2003, 17:49:30  
 Job time : 12.9612 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:00:31 ; Search time 6.31025 Seconds  
(without alignments)  
253.383 Million cell updates/sec

Title: US-09-848-834A-10  
Perfect score: 186  
Sequence: 1 FNNFTVSFWLRVPRKVSASHLEGPLHMSYGLRXP 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	112	60.2	1314	1	TEPX_CLOTE
2	63	33.9	431	1	PURA_LEGPN
3	61.5	33.1	92	1	GON1_RAT
4	61	32.8	1274	1	BXF_CLOBO
5	61	32.8	1295	1	BXG_CLOBO
6	59	31.7	1295	1	BXA1_CLOBO
7	58.5	31.5	1250	1	BXE_CLOBO
8	58	31.2	1290	1	BXB_CLOBO
9	57.5	30.9	90	1	GON1_MOUSE
10	57	30.6	1051	1	VP2_AHSV6
11	56.5	30.4	1250	1	BXE_CLOBU
12	56	30.1	1295	1	BXA2_CLOBO
13	54.5	29.3	92	1	GON1_HUMAN
14	54.5	29.3	1290	1	BXC1_CLOBO
15	54	29.0	67	1	GON1_MACMU
16	54	29.0	91	1	GON1_PIG
17	53.5	28.8	760	1	AMY_CLOAB
18	53.5	28.8	3122	1	DPROZ_MOUSE
19	52	28.0	61	1	GON1_SHEEP
20	52	28.0	63	1	GON1_MESAU
21	52	28.0	89	1	GON1_XENLA
22	52	28.0	92	1	GON1_TUPGB
23	52	28.0	449	1	VNSNS_INSVN
24	52	28.0	464	1	VNSNS_TSWV1
25	52	28.0	467	1	VNSNS_TSWV1
26	51.5	27.7	90	1	GON8_RANDY
27	51.5	27.7	521	1	UBP3_HUMAN
28	51	27.4	95	1	GON1_MORSA
29	51	27.4	292	1	CC22_ORYSA
30	51	27.4	1196	1	BXCN_CLOBO
31	50.5	27.2	537	1	COX1_SCHPO
32	50.5	27.2	1450	1	SREJ_STRPT
33	50	26.9	241	1	CRTA_RHOCA

34	50	26.9	842	1	ABC6_HUMAN
35	49.5	26.6	95	1	GON1_PAGMA
36	49.5	26.6	95	1	GON1_SPAPD
37	49.5	26.6	3130	1	DPROZ_HUMAN
38	49	26.3	523	1	C9B1_GLYEC
39	49	26.3	551	1	SYK_BRUME
40	49	26.3	551	1	SYK_BRUSU
41	48.5	26.1	92	1	GON1_CAVPO
42	48.5	26.1	5217	1	HTS1_COCCA
43	48	25.8	10	1	GON1_ALIMI
44	48	25.8	92	1	GON1_CHICK
45	48	25.8	94	1	GON1_HAPBU

## ALIGNMENTS

RESULT 1	ID	TEPX_CLOTE	STANDARD	PRT	1314 AA.
AC	P04958	TEPX_CLOTE			
DT	13-AUG-1987	(rel. 05, last sequence update)			
DT	13-AUG-1987	(rel. 05, last sequence update)			
DT	15-SEP-2003	(rel. 42, last annotation update)			
DE	Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains: Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy chain (Tetanus toxin chain H)].				
DE	Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy chain (Tetanus toxin chain H)].				
GN	TEPX OR CTP60.				
OS	Clostridium tetani.				
OC	Plasmid pE88, and Plasmid 75 kbp.				
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.				
OX	NCBI_TaxID=1513;				
RN	[1]	SEQUENCE FROM N.A.			
RP	PLASMID-75 kbp.				
RC	MEDLINE=87053814; PubMed=3536478;				
RX	Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J., Weller U., Hudek M., Habermann E., Niemann H.,				
RA	"Tetanus toxin: primary structure, expression in E. coli, and homology with botulinum toxins.";				
RT	EMBO J. 5:2495-2502(1986).				
RL	[2]	SEQUENCE FROM N.A.			
RP	STRAIN-CN3911; PLASMID-75 kbp.				
RC	MEDLINE=87040747; PubMed=3774547;				
RX	Faltweather N.F., Lyness V.A.;				
RA	"The complete nucleotide sequence of tetanus toxin.";				
RT	Nucleic Acids Res. 14:7809-7812(1986).				
RN	[3]	SEQUENCE FROM N.A.			
RP	STRAIN-Massachusetts / E88; PLASMID-pE88;				
RC	MEDLINE=22457253; PubMed=12552129;				
RX	Bruggemann H., Baumann S., Fricke W.F., Wietzer A., Liesegang H., Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A., Gottschalk G.;				
RA	"The genome sequence of Clostridium tetani, the causative agent of tetanus disease.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).				
RL	[4]	SEQUENCE OF 742-1314 FROM N.A.			
RP	PLASMID-75 kbp.				
RC	MEDLINE=86085672; PubMed=3510187;				
RX	Faltweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;				
RA	"Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in Escherichia coli.";				
RT	J. Bacteriol. 165:21-27(1986).				
RN	[5]	PARTIAL SEQUENCE, AND DISULFIDE BONDS.			
RP	MEDLINE=90201034; PubMed=2108021;				
RX	Kriegstein K., Henschen A., Weller U., Habermann E.;				
RA	"Arrangement of disulfide bridges and positions of sulfhydryl groups in tetanus toxin.";				
RT					

RL Eur. J. Biochem. 188:39-45(1990).  
 RN [6]  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE-92037649; PubMed-1935979;  
 RA Krieglstein K.G., Henschel A.H., Weller U., Habermann E.;  
 RT "Limited proteolysis of tetanus toxin. Relation to activity and  
 RL identification of cleavage sites."; Eur. J. Biochem. 202:41-51(1991).  
 RN [7]  
 RP IDENTIFICATION AS ZINC-PROTEASE.  
 RX MEDLINE-93010948; PubMed-1396558;  
 RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,  
 RA Montecucco C.;  
 RT "Tetanus toxin is a zinc protein and its inhibition of  
 RT neurotransmitter release and protease activity depend on zinc."; Eur. J. Biochem. 202:41-51(1991).  
 RL [8]  
 RN IDENTIFICATION OF SUBSTRATE.  
 RP MEDLINE-93063293; PubMed-1331807;  
 RX Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,  
 RA Dasgupta B.R., Montecucco C.;  
 RT "Tetanus and botulinum B neurotoxins block neurotransmitter release  
 RT by proteolytic cleavage of synaptobrevin."; Nature 359:832-835(1992).  
 RN [9]  
 RL X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.  
 RP MEDLINE-97475217; PubMed-9334741;  
 RX Dmland T.C., Wingerl L.M., Swaminathan S., Furey W.F., Schmidt J.J.,  
 RA Sax M.;  
 RT "Structure of the receptor binding fragment HC of tetanus  
 RT neurotoxin."; Nat. Struct. Biol. 4:788-792(1997).  
 CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES. IS INTERNALIZED  
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-1-PHE-77  
 CC BOND OF SYNAPTOSOMAL-2.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-1-Phe-77 bond in  
 CC synaptobrevin 2.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (by similarity).  
 CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO  
 CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE  
 CC AND ARE NON-TOXIC AFTER SEPARATION.  
 CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO  
 CC GANGLIOSIDE RECEPTORS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X04436; CAA28033.1; -;  
 DR EMBL; X06214; CAA29564.1; -;  
 DR EMBL; AF528097; AAO37454.1; -;  
 DR EMBL; M12739; AAA23282.1; -;  
 DR PIR; A25689; BTCLTN.  
 DR PDB; 1AF9; 29-APR-98.  
 DR PDB; 1ABD; 14-OCT-98.  
 DR PDB; 1DOH; 27-MAR-00.  
 DR PDB; 1DFQ; 24-MAR-00.  
 DR PDB; 1DIW; 24-MAR-00.  
 DR PDB; 1DLI; 24-MAR-00.  
 DR PDB; 1FV3; 05-SEP-01.  
 DR MEROPS; M27.001; -;  
 DR InterPro; IPR000395; Bontoxilysin.  
 DR InterPro; IPR006025; Zn\_Mtpeptidase.  
 DR Pfam; PF01742; Peptidase\_M27; 1.

DR PRINTS; PR00760; BONTOTOXILYSIN.  
 DR ProDom; PD001963; Bontoxilysin; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;  
 KW 3D-structure; Complete proteome.  
 FT INIT\_MET 0 0  
 FT CHAIN 1 456  
 FT METAL 457 1314  
 FT ACT\_SITE 232 232  
 FT METAL 233 233  
 FT TRANSMEM 236 236  
 FT TRANSMEM 226 246  
 FT DISULFID 669 689  
 FT DISULFID 438 466  
 FT DISULFID 1076 1092  
 FT HELIX 876 882  
 FT TURN 883 883  
 FT STRAND 884 891  
 FT TURN 892 893  
 FT STRAND 894 897  
 FT TURN 904 907  
 FT TURN 909 910  
 FT STRAND 912 915  
 FT STRAND 920 925  
 FT TURN 928 929  
 FT STRAND 932 935  
 FT HELIX 938 940  
 FT TURN 941 946  
 FT STRAND 949 956  
 FT HELIX 962 968  
 FT TURN 969 970  
 FT STRAND 972 977  
 FT STRAND 980 981  
 FT HELIX 983 985  
 FT STRAND 987 995  
 FT TURN 996 997  
 FT STRAND 998 1004  
 FT TURN 1006 1007  
 FT STRAND 1010 1016  
 FT STRAND 1020 1020  
 FT TURN 1021 1022  
 FT STRAND 1031 1037  
 FT TURN 1039 1040  
 FT STRAND 1042 1047  
 FT TURN 1048 1049  
 FT STRAND 1050 1056  
 FT TURN 1058 1059  
 FT STRAND 1068 1074  
 FT TURN 1079 1080  
 FT STRAND 1082 1091  
 FT HELIX 1097 1105  
 FT TURN 1106 1107  
 FT STRAND 1112 1112  
 FT STRAND 1114 1114  
 FT TURN 1116 1117  
 FT STRAND 1120 1120  
 FT TURN 1122 1122  
 FT STRAND 1123 1124  
 FT STRAND 1127 1131  
 FT HELIX 1132 1134  
 FT TURN 1135 1136  
 FT STRAND 1137 1141  
 FT TURN 1144 1145  
 FT STRAND 1148 1152  
 FT TURN 1155 1158  
 FT TURN 1159 1162  
 FT STRAND 1163 1166  
 FT STRAND 1173 1178  
 FT TURN 1184 1185  
 FT STRAND 1188 1188  
 FT STRAND 1190 1190  
 FT TURN 1191 1192  
 FT STRAND 1193 1201



FT	CHAIN	24	92	PROGMOADOLIBERIN I.
FT	PEPTIDE	24	33	CONADOLIBERIN I.
FT	PEPTIDE	37	92	PROACTININ RELEASE-INHIBITING FACTOR I.
FT	ACT_SITE	26	26	APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
FT	MOD_RES	24	24	PYROLIDONE CARBOXYLIC ACID.
FT	MOD_RES	33	33	AMIDATION (G-34 PROVIDE AMIDE GROUP).
SO	SEQUENCE	92 AA;	10500 MW;	49AB5C64DAB3B3 CR664;

  

Query Match	Best Local Similarity	33.1%;	Score 61.5;	DB 1;	Length 92;
Matches 15;	Conservative 3;	Pred. No. 0.078;	Mismatches 4;	Indels 7;	Gaps 2;

  

Db	12	VPKVSAS-----HLEG-PSLHMSYGLRP	33
	::: :		
	4	IPKILAAVLLTVCLEGGSSQHSWSGLRP	32

  

RESULT 4
BXF_CLOBO
ID BXF_CLOBO
AC P30996;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BONT/F)
DE (bontoxilysin F).
GN BONTF.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 23387;
RX MEDLINE=93012902; PubMed=1398040;
RA East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA Roberts T.A., Thompson D.E.,
RT "Sequence of the gene encoding type F neurotoxin of Clostridium
RT botulinum."
RT FEWS Microbiol. Lett. 75:225-230(1992).
RL
[2]
RN SEQUENCE OF 1-64 FROM N.A.
RP STRAIN=Hobbs FT10;
RX MEDLINE=94297488; PubMed=7764998;
RA East A.K., Collins M.D.;
RT "Conserved structure of genes encoding components of botulinum
RT neurotoxin complex M and the sequence of the gene coding for the
RT nontoxic component in nonproteolytic Clostridium botulinum type F."
RT Curr. Microbiol. 29:69-77(1994).
RL
[3]
RN SEQUENCE OF 634-1002 FROM N.A.
RP MEDLINE=94013372; PubMed=8408542;
RX Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F."
RT J. Clin. Microbiol. 31:2255-2262(1993).
RL
[4]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94230352; PubMed=8175689;
RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA Roques B., Eykhe E.M., Suehno T.C., Jain R., Niemann H.;
RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT F botulinum neurotoxins and tetanus toxin."
RT J. Biol. Chem. 269:12764-12772(1994).
RL
-I- PUNCTION. BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 56-GLN- -LYS-59
CC BOND OF SYNAPTOSOMAL-1 AND -2.
CC -I- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the

```

CC      neurocytosis apparatus, synaptobrevis, SNRP25 or syntaxin. NO
CC      detected action on small molecule substrates.
CC      -1- COPACITOR: binds 1 zinc ion per subunit (By similarity).
CC      -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC      HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC      WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC      FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC      BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC      or send an email to license@isb-slb.ch).
CC      -----
DR      EMBL; M92906; AAA23263.1; -.
DR      EMBL; S73676; AAC60475.1; -.
DR      EMBL; X70820; CAA50151.1; -.
DR      EMBL; X70816; CAA50147.1; -.
DR      PIR; I40813; I40813.
DR      PIR; S48109; S48109.
DR      HSSP; P10845; 3B7A.
DR      MEROPS; M27.002; -.
DR      InterPro; IPR000395; Bontoxilysin.
DR      InterPro; IPR006025; Zn_MTPeptide.
DR      Pfam; PF01742; Peptidase_M27; 1.
DR      ProDom; PD001963; Bontoxilysin; 1.
DR      PROSITE; PS00142; ZINC_PROTEASE; 1.
DR      Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; zinc.
DR      EMBL; CHAIN 1 436 BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
DR      EMBL; CHAIN 1 436 BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.
ET      METAL 228 227 ZINC (CATALYTIC) (BY SIMILARITY).
ET      ACTL SITE 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
ET      METAL 231 231 ZINC (CATALYTIC) (BY SIMILARITY).
ET      DISULFID 429 445 INTERCHAIN (PROBABLE).
SQ      SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;

Query Match      32.8%; Score 61; DB 1; Length 1274;
Best Local Similarity 57.1%; Pred. No. 1.5;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 FNNFTVSFWLRYPK 14
      1 :|||||:|
DB      930 YQNESISFWRIK 943

RESULT 5
BXG_CLOBO      STANDARD:      PRT: 1296 AA.
AC      Q60393;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BONT/G)
DE      (Bontoxilysin G).
GN      BontG.
OS      Clostridium botulinum.
OC      Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC      Clostridium.
OX      NCBI_TaxID=1491;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=13 / 30;
RX      MEDLINE=94092745; PubMed=8268233;
RA      Campbell K., Collins M.D., East A.K.;
RT      "Nucleotide sequence of the gene coding for Clostridium botulinum
RT      (Clostridium argentinense) type G neurotoxin: genealogical comparison
RT      with other clostridial neurotoxins";

```



RT and purification of two tryptic fragments. Proteolytic action near  
RT the COOH-terminus of the heavy subunit destroys toxin-binding  
RT activity.".  
RL Eur. J. Biochem. 151:75-82(1985).  
RN [11]  
RP IDENTIFICATION OF SUBSTRATE.  
RX MEDLINE=94063091; PubMed=8243676;  
RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,  
Bentinati F., Wilson M.C., Montecucco C.,  
RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct  
RT COOH-terminal peptide bonds".  
RL FEBS Lett. 335:99-103(1993).  
RN [12]  
RP IDENTIFICATION OF SUBSTRATE.  
RX MEDLINE=94124495; PubMed=8294407;  
RA Binz T., Blas J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,  
RA Jahn R., Niemann H.,  
RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins".  
RL J. Biol. Chem. 269:1617-1620(1994).  
RN [13]  
RP MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.  
RX MEDLINE=21556941; PubMed=11700044;  
RA Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.,  
RT "Site-directed mutagenesis identifies active-site residues of the  
RT light chain of botulinum neurotoxin type a".  
RN Biochem. Biophys. Res. Commun. 288:1231-1237(2001).  
RL [14]  
RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).  
RX MEDLINE=98455071; PubMed=9783750;  
RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.,  
RT "Crystal structure of botulinum neurotoxin type A and implications  
RT for toxicity".  
RL Nat. Struct. Biol. 5:898-902(1998).  
CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin  
CC binds with high affinity to peripheral neuronal presynaptic  
CC membrane, is then internalized by receptor-mediated endocytosis.  
CC The C-terminus of the heavy chain (H) is responsible for the  
CC adherence of the toxin to the cell surface while the N-terminus  
CC mediates transport of the light chain from the endocytic vesicle  
CC to the cytosol. After translocation, the light chain (L)  
CC hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking  
CC neurotransmitter release. Inhibition of acetylcholine release  
CC results in flaccid paralysis, with frequent heart or respiratory  
CC failure.  
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No  
CC detected action on small molecule substrates.  
CC -1- COFACTOR: Binds 1 zinc ion per subunit.  
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a  
CC heavy chain (H).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for  
CC the treatment of strabismus and blepharospasm associated with  
CC dystonia and cervical dystonia. Also used for the treatment of  
CC hemifacial spasm and a number of other neurological disorders  
CC characterized by abnormal muscle contraction.  
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of  
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
CC -1- DATABASE: NAME-BOTOX product information Web site:  
CC WWW="http://www.botox.com/index.jsp#peptideinfo".  
CC -1- DATABASE: NAME-Protein Spotlight;  
CC NOTE=Issue 19 of February 2002;  
CC WWW="http://www.expasy.org/spotlight/articles/spilt019.html".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----

DR EMBL: X52066; CAA36289.1; -  
DR EMBL: M30196; AAA23262.1; -  
DR EMBL: X92973; CAA63551.1; -  
DR EMBL: D67030; BAA11051.1; -  
DR EMBL: M27892; AAA23269.1; -  
DR PIR: A35294; BTCLAB.  
DR PDB: 3BTA; 01-OCT-99.  
DR MEROPS: M27.002; -  
DR InterPro: IPR000395; Bontoxilysin.  
DR InterPro: IPR006025; Zn\_MTPeptide.  
DR Pfam: PF01742; Peptidase\_M27; 1.  
DR ProDom: PD001963; Bontoxilysin; 1.  
DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;  
KW Pharmaceutical; 3D-structure.  
FT INTR MET 0  
FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.  
FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.  
FT METAL 222 222 ZINC (CATALYTIC).  
FT ACT\_SITE 223 223 ZINC (CATALYTIC).  
FT METAL 226 226 ZINC (CATALYTIC).  
FT METAL 261 261 ZINC (CATALYTIC).  
FT DISULFID 429 453 INTERCHAIN.  
FT DISULFID 1234 1279 POTENTIAL.  
FT TRANSMEM 626 646 POTENTIAL.  
FT TRANSMEM 655 675 V -> A.  
FT VARIANT 26 26 E->A: DRASTIC DECREASE IN ENZYMATIC  
FT MUTAGEN 261 261 ACTIVITY.  
FT MUTAGEN 265 265 F->A: DECREASE IN ENZYMATIC ACTIVITY.  
FT MUTAGEN 365 365 Y->A: DECREASE IN ENZYMATIC ACTIVITY.  
FT CONFLICT 1 1 Y -> Q (IN REF. 1).  
FT CONFLICT 479 479 E -> P (IN REF. 9).  
FT CONFLICT 875 875 T -> L (IN REF. 8).  
FT CONFLICT 891 891 S -> K (IN REF. 8).  
SQ SEQUENCE 1295 AA; 149322 MW; 858342F754862579 CRC64;  
Query Match 31.7%; Score 59; DB 1; Length 1295;  
Best Local Similarity 57.1%; Pred. No. 2.9;  
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
OY 1 FNNFTVSFWLRPK 14  
Db 937 YENFTSEWIRPK 950  
: 11: 111:111  
RESULT 7  
BXL\_CLOBO STANDARD; PRT; 1250 AA.  
AC 000496;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)  
DE (Bontoxilysin E).  
OS Clostridium botulinum.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1491;  
RN [11]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Beluga;  
RX MEDLINE=92181428; PubMed=1543481;  
RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;  
RT "Sequences of the botulinum neurotoxin E derived from Clostridium  
RT ATCC 43181 and ATCC 43755".  
RL Biochem. Biophys. Res. Commun. 183:107-113(1992).  
RN [12]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92174922; PubMed=1541280;  
RA Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;  
RT "The complete amino acid sequence of the Clostridium botulinum type-E



RT "Botulinum neurotoxin type B (strain 657): partial sequence and  
RT similarity with tetanus toxin.";  
RL Biochimie 70:811-817(1988).  
RN [5]  
RP SEQUENCE OF 1-16 AND 441-458.  
RC STRAIN-OKRA: Pubmed=3888113;  
RA MEDLINE=85197963; Dasgupta B.R.;  
RA Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;  
RT "Partial amino acid sequences of botulinum neurotoxins types B and  
RT E.";  
RL Arch. Biochem. Biophys. 238:544-548(1985).  
RN [6]  
RP IDENTIFICATION AS ZINC-PROTEASE.  
RA MEDLINE=93054694; Pubmed=1429690;  
RA Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;  
RT "Botulinum neurotoxins are zinc proteins.";  
RL J. Biol. Chem. 267:23479-23483(1992).  
RN [7]  
RP IDENTIFICATION OF SUBSTRATE.  
RX MEDLINE=93063293; Pubmed=1331807;  
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,  
RA Dasgupta B.R., Montecucco C.;  
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release  
RT by proteolytic cleavage of synaptobrevin.";  
RL Nature 359:832-835(1992).  
CC -I- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-1-PHE-77 BOND OF  
CC SYNAPTOBREVIN-2.  
CC -I- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO  
CC detected action on small molecule substrates.  
CC -I- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -I- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A  
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,  
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL  
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.  
CC -I- SUBCELLULAR LOCATION: Secreted.  
CC -I- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF  
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.  
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M81186; AAA23211.1; -;  
DR EMBL: Z11934; CAAT7991.1; -;  
DR EMBL: X70817; CA50148.1; -;  
DR PIR: A48940; A48940.  
DR PDB: 1EPW; 01-NOV-00.  
DR PDB: 1F31; 01-NOV-00.  
DR PDB: 1F82; 16-AUG-00.  
DR PDB: 1F83; 16-AUG-00.  
DR PDB: 1F0H; 06-DEC-00.  
DR PDB: 1G9A; 13-NOV-02.  
DR PDB: 1G9B; 13-NOV-02.  
DR PDB: 1G9C; 13-NOV-02.  
DR PDB: 1G9D; 13-NOV-02.  
DR PDB: 1IIE; 21-NOV-01.  
DR MEROPS: M27.002; -;  
DR InterPro: IPR000395; Bontoxilysin.  
DR InterPro: IPR006025; Zn\_MTPpeptidase.  
DR Pfam: PF01742; Peptidase\_M27.1.  
DR PRINTS: PR00760; BONTOXILYSIN.  
DR ProDom: PD001963; Bontoxilysin; 1.

DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;  
KW 3D-structure.  
FT INT\_MET 0 0  
FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.  
FT METAL 441 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.  
FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 230 230 BY SIMILARITY.  
FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).  
FT DISULFID 436 445 INTERCHAIN (PROBABLE).  
FT CONFLICT 29 29 T -> M (IN REF. 4).  
FT CONFLICT 217 217 R -> G (IN REF. 2).  
FT CONFLICT 224 224 A -> S (IN REF. 2).  
FT CONFLICT 463 463 S -> R (IN REF. 4).  
SQ SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;  
  
Query Match 31.2%; Score 58; DB 1; Length 1290;  
Best Local Similarity 64.3%; Pred. No. 4.1;  
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 FNNFVSFWLRVPK 14  
Db 922 FLDFSVSWIRIPK 935  
  
RESULT 9  
GONI\_MOUSE  
ID GONI\_MOUSE STANDARD; PRT; 90 AA.  
AC P13562; 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LH-RH I)  
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing  
DE hormone I) (GNRH I) (Laliberin I); Prolactin release-inhibiting factor  
DE I].  
GN GNRI1 OR GNRI.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87069928; Pubmed=3024317;  
RA Mason A.J., Haylick J.S., Zoeller R.T., Young W.S. III,  
RA Phillips H.S., Nikolic K., Seeburg P.H.;  
RT "A deletion truncating the gonadotropin-releasing hormone gene is  
RT responsible for hypogonadism in the hpg mouse.";  
CC Science 234:1366-1371(1986).  
CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS, IT STIMULATES  
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
CC HORMONES.  
CC -I- SUBCELLULAR LOCATION: Secreted.  
CC -I- SIMILARITY: Belongs to the GNRH family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M14872; AAA37717.1; -;  
DR PIR: A47578; RHMSG.  
DR MGD: MGI:95789; GnRH.  
DR InterPro: IPR002012; GNRH.  
DR InterPro: IPR004079; GonadolibirinI.  
DR Pfam: PF00446; GnRH.1.  
DR PRINTS: PR01541; GONADOLIBRINI.  
DR PROSITE: PS00473; GNRH.1.  
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
KW Placenta; Signal; Pyroglutamate carboxylic acid.

```

FT SIGNAL 1 21
FT CHAIN 22 90
FT PEPTIDE 22 31
FT ACT_SITE 24 24
FT MOD_RES 22 22
FT MOD_RES 31 31
SQ SEQUENCE 90 AA: 10337 MW: 10766FA826EAD9 CRC64;

Query Match
Best local Similarity 30.9%; Score 57.5; DB 1; Length 90;
Matches 12; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 20 LEG-PSLHWSYGLRP 33
DB 16 LEGCSQHSYGLRP 30

RESULT 10
VP2_AHSV6 STANDARD; PRT; 1051 AA.
AC 071024;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Outer capsid protein VP2.
GN S2 OR I2.
OS African horse sickness virus 6 (AHSV-6) (African horse sickness virus
OS (serotype 6)).
OC Viruses; dsRNA viruses; Reoviridae; Orbiviruses.
OX NCBI_Taxid=86060;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98278331; PubMed=9617769;
RA Williams C.F., Inoue T., Lucus A.M., Zanotto P., Roy P.;
RT "The complete sequence of four major structural proteins of African
RT horse sickness virus serotype 6: evolutionary relationships within
RT and between the orbiviruses.";
RL Virus Res. 53:53-73(1998).
CC -1- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5)
CC WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE
CC MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
CC -1- SIMILARITY: BELONGS TO THE REOVIRUSES VP2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF021235; AAC40994.1; -
CC InterPro; IPR001742; Orbl_VP2.
CC Pfam; PF00898; Orbl_VP2; 1.
CC Coat protein.
SQ SEQUENCE 1051 AA: 122326 MW: 2B04DB9E389F4B5F CRC64;

Query Match
Best local Similarity 47.6%; Score 57; DB 1; Length 1051;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 FNNFTVSEWLRVPKVSASHLE 21
DB 636 FSKRFVSYWYRVEKITTKHLE 656

RESULT 11
BXE_CLOBU STANDARD; PRT; 1250 AA.
ID BXE_CLOBU
AC P30995;
DT 01-JUL-1993 (Rel. 26, Created)

```

```

DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (Bont/E)
DE (Bontoxilysin E).
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_Taxid=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43181, and ATCC 43755;
RX MEDLINE=92181428; PubMed=1543481;
RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT "Sequences of the botulinum neurotoxin E derived from Clostridium
RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT ATCC 43181 and ATCC 43755).";
RL J. Gen. Microbiol. 137:519-525(1991).
RN [2]
RP SEQUENCE OF 1-251 FROM N.A.
RC STRAIN=BL6340;
RX MEDLINE=91237316; PubMed=2033376;
RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
RA Yokosawa N., Yashiki T., Oguma K.;
RT "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
RT type E toxin gene from Clostridium butyricum strain BL6340.";
RL J. Gen. Microbiol. 137:519-525(1991).
RN [3]
RP SEQUENCE OF 1-48.
RC STRAIN=5262;
RA Gimenez J., Foley J., Dasgupta B.R.;
RT Neurotoxin type E from Clostridium botulinum and C. butyricum;
RT partial sequence and comparison.";
RL FASEB J. 2:41750-41750(1988).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
CC detected action on small molecule substrates.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (by similarity).
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X62088; CAA43998.1; -
CC EMBL; X53180; CAA37321.1; -
CC PIR; JH0256; JH0256.
CC HSSP; P10845; 3BTA.
CC MEROPS; M27.002; -.
CC InterPro; IPR000395; Bontoxilysin.
CC InterPro; IPR006025; Zn_MTPeptide.
CC Pfam; PF01742; Peptidase_M27; 1.
CC PRINTS; PR00760; BONTOTOXILYSIN.
CC PRODOM; PD001963; Bontoxilysin; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Neurotoxin; transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0

```

```

FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 212 212 BY SIMILARITY.
FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 411 425 INTERCHAIN (PROBABLE).
FT CONFLICT 229 229 R -> M (IN REF. 2).
SQ SEQUENCE 1250 AA; 143265 MW; 8171B582C312857 CRC64;

Query Match 30.4%; Score 56.5; DB 1; Length 1250;
Best Local Similarity 22.4%; Pred. No. 6.3;
Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;

OY 1 FNNFTVSFMRVP-----KVSASHLE-----GP 23
   ||:|||||:|
Db 911 YKNFSISFWRIPYNDKIVNNEYTIICNRDNNSGMKVSLNNEIITWLTQDSGLNQ 970
   ||:|

OY 24 SLHMSYG 30
   |::|
Db 971 KLAFTNYG 977

RESULT 12
BXA2_CLOBO STANDARD: PRT; 1295 AA.
ID BXA2_CLOBO
AC Q45894; P77780;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)
DE (Bontoxilysin A) (BOTOX) [contains: Botulinum neurotoxin A, light-chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kyoto-F;
RA MEDLINE=94143603; PubMed=8310180;
RA Willem A., East A.K., Lawson P.A., Collins M.D.;
RT "Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A associated with infant botulism: comparison with other clostridial neurotoxins.";
RL Res. Microbiol. 144:547-556(1993).
RN [2]
RP SEQUENCE OE 1-65 FROM N.A.
RC STRAIN=Kyoto-F;
RA MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding components of the botulinum toxin complex in proteolytic Clostridium botulinum types A, B, and F: evidence of chimeric sequences in the gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC -I- FUNCTION: Inhibits acetylcholine release. The botulinum toxin binds with high affinity to peripheral neuronal presynaptic membrane, is then internalized by receptor-mediated endocytosis. The C-terminus of the heavy chain (H) is responsible for the adherence of the toxin to the cell surface while the N-terminus mediates transport of the light chain from the endocytic vesicle to the cytosol. After translocation, the light chain (L) hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking neurotransmitter release. Inhibition of acetylcholine release results in flaccid paralysis, with frequent heart or respiratory failure (By similarity).
CC -I- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO detected action on small molecule substrates.
CC -I- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a heavy chain (H) (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted.

```

```

CC -I- MISCELLANEOUS: There are seven antigenically distinct forms of botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X73423; CAA51824.1; -
DR EMBL: X87974; CAA61234.1; -
DR PIR: I40645; I40645.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR006025; Zn_MpPeptidase.
DR Pfam: PF01742; Peptidase_M27; 1.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; FALSE_NEG.
KM Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0
FT CHAIN 0 447 BY SIMILARITY.
FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 223 223 BY SIMILARITY.
FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 429 453 INTERCHAIN (BY SIMILARITY).
FT DISULFID 1234 1279 BY SIMILARITY.
FT TRANSMEM 626 646 POTENTIAL.
FT TRANSMEM 655 675 POTENTIAL.
SQ SEQUENCE 1295 AA; 149279 MW; 5DA04A13D98D6372 CRC64;

Query Match 30.1%; Score 56; DB 1; Length 1295;
Best Local Similarity 50.0%; Pred. No. 7.7;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 FNNFTVSFMRVP 14
   ||:|||||
Db 937 YENFSTFWIKIPK 950

RESULT 13
GONL_HUMAN STANDARD: PRT; 92 AA.
ID GONL_HUMAN
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progredinoliberin I precursor [Contains: Gonadoliberin I (LH-RH I) (Luteinizing hormone-releasing hormone I) (Gonadorelin); GnRH-associated peptide I].
DE hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated peptide I).
GN GNRI1 OR GNRI OR LHRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=8936682; PubMed=2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing hormone gene.";
RL Nucleic Acids Res. 17:6403-6403(1989).
RN [2]
RP SEQUENCE FROM N.A.
RP MEDLINE=86094336; PubMed=2867548;
RA Adelman J.P., Mason A.O., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting

```

RT factor in human and rat.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANT SER-16.  
 RX MEDLINE=85012739; PubMed=6090951;  
 RA Seeburg P.H., Adelman J.P.;  
 RT "Characterization of cDNA for precursor of human luteinizing hormone  
 RT releasing hormone.";  
 RL Nature 311:666-668(1984).  
 RN [4]  
 RP SEQUENCE OF 24-33.  
 RX MEDLINE=83126573; PubMed=6760865;  
 RA Tan L., Rousseau P.;  
 RT "The chemical identity of the immunoreactive LHRH-like peptide  
 RT biosynthesized in the human placenta.";  
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).  
 RN [5]  
 RP VARIANT SER-16.  
 RX MEDLINE=99318093; PubMed=10391209;  
 RA Carilli M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [6]  
 RP ERRATUM.  
 RA Carilli M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RT Nat. Genet. 23:373-373(1999).  
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES  
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING  
 CC HORMONES.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),  
 CC Lutrephase or Lutrelief (Ferring Pharmaceuticals) and Kelsorm  
 CC (Seono).  
 CC -1- SIMILARITY: Belongs to the GNRH family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X01059; CAA25526.1; -  
 DR EMBL: M12578; AAA55916.1; -  
 DR EMBL: X15215; CAA33285.1; -  
 DR PIR: S05308; RHUOG.  
 DR Genew: HGNC:4419; GNRH1.  
 DR MIM: 152760; -  
 DR GO: GO:0005625; C:soluble fraction; TAS.  
 DR GO: GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.  
 DR GO: GO:0007267; P:cell-cell signaling; TAS.  
 DR GO: GO:0007275; P:development; TAS.  
 DR GO: GO:0008285; P:negative regulation of cell proliferation; TAS.  
 DR GO: GO:0007165; P:signal transduction; TAS.  
 DR InterPro: IPR02012; GNRH.  
 DR InterPro: IPR004079; Gonadolibertin.  
 DR Pfam: PF00446; GNRH; 1.  
 DR PRINTS: PRO1541; GONADOLIBERIN.  
 DR PROSITE: PS00473; GNRH; 1.  
 DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;  
 KW Placenta; Pharmaceutical; Signal; Polymorphism;  
 KW Pyroglutamate; Carboxylic acid.  
 FT SIGNAL 1 23  
 FT CHAIN 24 92 PROGONADOLIBERIN I.  
 FT PEPTIDE 24 33 GONADOLIBERIN I.  
 CC

FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.  
 FT ACT\_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL  
 FT MOD\_RES 24 24 ACTIVITY.  
 FT MOD\_RES 33 33 PYROLIDONE CARBOXYLIC ACID.  
 FT VARIANT 16 16 AMIDATION (G-34 PROVIDE AMIDE GROUP).  
 FT W -> S (IN OBSNP:6185).  
 FT /FTID=VAR\_013943.  
 SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;  
 Query Match 29.3%; Score 54.5; DB 1; Length 92;  
 Best Local Similarity 73.3%; Pred. No. 0.74;  
 Matches 11; Conservative 1; Mismatches 2; Indels 1; Gaps 1;  
 QY 20 LGS-PSLHWSYGLRP 33  
 DB 18 VEGCSQSHWSYGLRP 32  
 RESULT 14  
 ID BXCL\_CLOBO STANDARD; PRT; 1290 AA.  
 AC P18640;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BONT/C1)  
 DE (Bontoxilysin C1).  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxId=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90370487; PubMed=2204031;  
 RA Hausner D., Eklund M.W., Kurazono H., Binz T., Niemann H., Gill D.M.,  
 RA Boquet P., Popoff M.R.;  
 RT "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";  
 RL Nucleic Acids Res. 18:4924-4924(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Type C Stockholm / C-ST;  
 RX MEDLINE=91024998; PubMed=2222445;  
 RA Kimura K., Fujii N., Suzuki K., Murakami T., Indoh T.,  
 RA Yokosawa N., Takeshi K., Syuto B., Oguma K.;  
 RT "The complete nucleotide sequence of the gene coding for botulinum  
 RT type C1 toxin in the C-ST phage genome.";  
 RL Biochem. Biophys. Res. Commun. 171:1304-1311(1990).  
 RN [3]  
 RP SEQUENCE OF 2-25.  
 RC STRAIN-Type C Stockholm / C-ST;  
 RX MEDLINE=88153072; PubMed=2450068;  
 RA Tsuzuki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,  
 RA Oguma K.;  
 RT "Establishment of a monoclonal antibody recognizing an antigenic site  
 RT common to Clostridium botulinum type B, C1, D, and E toxins and  
 RT tetanus toxin.";  
 RL Infect. Immun. 56:898-902(1988).  
 RN [4]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=94038966; PubMed=7901002;  
 RA Blas J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.;  
 RT "Botulinum neurotoxin C1 blocks neurotransmitter release by means of  
 RT cleaving HPC-1/syntaxin.";  
 RL EMBO J. 12:4821-4828(1993).  
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 CC ENDOPEPTIDASE THAT CLEAVES SYNTAXIN.  
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No  
 CC detected action on small molecule substrates.  
 CC



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 17:29:11 : Search time 32.964 Seconds  
(without alignments)  
266.163 Million cell updates/sec

Title: US-09-848-834A-10

Perfect score: 186  
Sequence: 1 FNNFTVSFWLRVPKVSASHLEBPSLHWSYGLRPX 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	60.2	451	2	09LA13
2	112	60.2	1310	2	093N27
3	62	33.3	1268	2	045851
4	61	32.8	1278	2	057236
5	58	31.2	361	2	045848
6	58	31.2	361	2	045846
7	58	31.2	441	2	09X708
8	58	31.2	1291	2	09ZAU8
9	58	31.2	1291	2	093G71
10	58	31.2	1291	2	0933K0
11	58	31.2	1291	2	008077
12	58	31.2	1291	2	08GR96
13	57	30.6	430	2	09XAV1
14	57	30.6	502	16	09X8T8
15	56.5	30.4	90	13	090Y63
16	56.5	30.4	367	2	045861

17	56.5	30.4	367	2	045862	045862 clostridium
18	56.5	30.4	1252	2	08KZM3	08KZM3 clostridium
19	56.5	30.4	1255	2	09PAR6	09PAR6 clostridium
20	56	30.1	804	10	08L853	08L853 arabidopsis
21	56	30.1	812	10	064620	064620 arabidopsis
22	55	30.1	1251	2	09K395	09K395 clostridium
23	55	29.6	1280	2	09ZAJ5	09ZAJ5 clostridium
24	54.5	29.3	476	16	08PM68	08PM68 xanthomonas
25	54.5	29.3	1285	2	045967	045967 clostridium
26	54.5	29.3	1285	2	09LBR1	09LBR1 clostridium
27	54.5	29.3	1291	2	093HT3	093HT3 clostridium
28	53	28.5	91	13	09PRH0	09PRH0 anguilla ja
29	53	28.5	606	16	08XIK6	08XIK6 clostridium
30	52.5	28.2	91	13	08L107	08L107 oryzias lat
31	52	28.0	466	12	08JXJ9	08JXJ9 tomato spot
32	52	28.0	466	12	08UXK0	08UXK0 tomato spot
33	52	28.0	467	12	08JXK2	08JXK2 tomato spot
34	52	28.0	467	12	037369	037369 tomato spot
35	52	28.0	467	12	037367	037367 tomato spot
36	52	28.0	467	12	08JVL0	08JVL0 tomato spot
37	52	28.0	467	12	08JXK4	08JXK4 tomato spot
38	51.5	27.7	91	13	09DGC8	09DGC8 oryzias lat
39	51.5	27.7	317	4	09Y2R8	09Y2R8 homo sapien
40	51.5	27.7	345	2	08VRF4	08VRF4 bradyrhizob
41	51.5	27.7	381	2	08VKZ1	08VKZ1 enterobacte
42	51.5	27.7	520	4	08WVDD	08WVDD homo sapien
43	51.5	27.7	1127	16	08G815	08G815 bifidobacte
44	51	27.4	130	2	093PY5	093PY5 pseudomonas
45	51	27.4	220	10	09C592	09C592 arabidopsis

## ALIGNMENTS

## RESULT 1

ID	09LA13	PRELIMINARY;	PRT;	451 AA.
AC	09LA13:			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
OS	Tetanus toxin (Fragment).			
OS	Clostridium tetani.			
OC	Bacteria: Firmicutes; Clostridia; Clostridiales; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1513;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=20886;			
RA	He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;			
RT	"Fragment C of Tetanus toxin."			
RL	Submitted (May-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF154828; AAF73267.1; -.			
DR	HSSP; P04958; 1ABD.			
DR	InterPro; IPR001064; Crystallin.			
DR	PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.			
FT	NON_TER			
FT	SEQUENCE			
FT	451 AA; 51823 MW; 69A8C5F030B6CD8E CRC64;			
Query Match		60.2%;	Score 112;	DB 2;
Best local similarity		100.0%;	Pred. NO. 1.2e-07;	Length 451;
Matches	21;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;
QY	1 FNNFTVSFWLRVPKVSASHLE 21			
DB	83 FNNFTVSFWLRVPKVSASHLE 103			
RESULT 2				
ID	093N27	PRELIMINARY;	PRT;	1310 AA.
AC	093N27:			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			

DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Tetanus toxin (TREMblrel. 23, Last annotation update)  
 OS Clostridium tetani.  
 CC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 CC Clostridium.  
 RN NCBI\_TaxID=1513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shumin Z., Dianliang L.;  
 RT "Cloning and sequence analysis of tetanus toxin gene."  
 RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF389424; AAK72964.2; -  
 DR InterPro: IPR000395; Bontoxilysin.  
 DR InterPro: IPR001064; Crystallin.  
 DR InterPro: IPR006025; Zn\_Mrpeptidse.  
 DR Pfam: PF01742; Peptidase\_M27; 1.  
 DR PRINTS: PR00760; BONTOXILYSIN.  
 DR ProDom: PD001963; Bontoxilysin; 1.  
 DR PROSITE: PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 1310 1310  
 SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;

Query Match 60.2%; Score 112; DB 2; Length 1310;  
 Best Local Similarity 100.0%; Pred. No. 4e-07;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FNNFTVSFWLRVPRVSASHLE 21  
 Db 948 FNNFTVSFWLRVPRVSASHLE 968

RESULT 3  
 ID 045851 PRELIMINARY; PRT; 1268 AA.  
 AC 045851;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Neurotoxin type F.  
 GN BONT/F.  
 OS Clostridium baratti.  
 CC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 CC Clostridium.  
 RN NCBI\_TaxID=1561;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-93252228; PubMed-8486245;  
 RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,  
 RA Richardson P.T.;  
 RT "Nucleotide sequence of the gene coding for Clostridium baratti type F  
 RT neurotoxin: Comparison with other clostridial neurotoxins."  
 RL FEMS Microbiol. Lett. 108:175-182(1993).  
 DR EMBL: X68262; CAA48329.1; -  
 DR HSSP: P10845; 3BTA.  
 DR MEROPS: M27\_002; -  
 DR InterPro: IPR000395; Bontoxilysin.  
 DR InterPro: IPR006025; Zn\_Mrpeptidse.  
 DR Pfam: PF01742; Peptidase\_M27; 1.  
 DR ProDom: PD001963; Bontoxilysin; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 SQ SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match 33.3%; Score 62; DB 2; Length 1268;  
 Best Local Similarity 64.3%; Pred. No. 5.3;  
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FNNFTVSFWLRVPR 14  
 Db 922 YONFSISFWVRIPK 935

RESULT 4  
 ID 057236 PRELIMINARY; PRT; 1278 AA.  
 AC 057236; 045863;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Botulinum neurotoxin type F (BONT/F protein).  
 GN BONT/F.  
 OS Clostridium botulinum.  
 CC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 CC Clostridium.  
 RN NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NCTC 10281;  
 RA Hutson R.A., Collins M.D.;  
 RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Elmore M.J., Bodevorth N.J., Whelan S.M., Minton N.P.;  
 RL Submitted (AUG-1994) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 635-1000 FROM N.A.  
 RC STRAIN-NCTC 1028;  
 RX MEDLINE-94013372; PubMed-8408542;  
 RA Campbell K., East A.K., Collins M.D.;  
 RT "Gene probes for identification of the botulin neurotoxin gene and  
 RT specific identification of neurotoxin types B, E, and F."  
 RL J. Clin. Microbiol. 31:2255-2262(1993).  
 RN [4]  
 RP SEQUENCE OF 1-27 FROM N.A.  
 RC STRAIN-LANGELAND;  
 RX MEDLINE-98404102; PubMed-9732534;  
 RA East A.K., Bhandari M., Hielm S., Collins M.D.;  
 RT "Analysis of the botulinum neurotoxin type F gene clusters in  
 RT proteolytic and nonproteolytic Clostridium botulinum and Clostridium  
 RT baratti."  
 RL Curr. Microbiol. 37:262-268(1998).  
 DR EMBL: X81714; CAA57358.1; -  
 DR EMBL: L35496; AAA23210.1; -  
 DR EMBL: X70821; CAA50152.1; -  
 DR EMBL: X99064; CAA67512.1; -  
 DR HSSP: P10845; 3BTA.  
 DR MEROPS: M27\_002; -  
 DR InterPro: IPR000395; Bontoxilysin.  
 DR InterPro: IPR006025; Zn\_Mrpeptidse.  
 DR Pfam: PF01742; Peptidase\_M27; 1.  
 DR PRINTS: PR00760; BONTOXILYSIN.  
 DR ProDom: PD001963; Bontoxilysin; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KW Neurotoxin.  
 SQ SEQUENCE 1278 AA; 147073 MW; A1BE1318431D6918 CRC64;

Query Match 32.8%; Score 61; DB 2; Length 1278;  
 Best Local Similarity 57.1%; Pred. No. 7.5;  
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FNNFTVSFWLRVPR 14  
 Db 931 YONFSISFWVRIPK 944

RESULT 5  
 ID 045848 PRELIMINARY; PRT; 361 AA.  
 AC 045848;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE Botulinum neurotoxin type B (Fragment).  
 GN BONT/B.

[illegible]

DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Neurotoxin type B.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1436;  
 RA Kitama N., Ferreira J.L., Baumstark B.R.;  
 RT "Characterization of six type A strains of Clostridium botulinum that  
 contain type B toxin gene sequences.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF295926; AK97137.1; -  
 DR InterPro: IPR000395; Bontoxilysin.  
 DR InterPro: IPR006025; Zn\_MTPeptide.  
 DR Pfam: PF01742; Peptidase\_M27; 1.  
 DR PRINTS: PR00760; BONTOTOXILYSIN.  
 DR ProDom: PD001963; Bontoxilysin; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 SQ SEQUENCE 1291 AA; 150824 MW; D7CA07BAE2EB8CD2 CRC64;

Query Match 31.2% Score 58; DB 2; Length 1291;  
 Best Local Similarity 64.3%; Pred. No. 20;  
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 FNNFTVSFWLRVPR 14  
 I:|:|||||:|:|  
 DB 923 FLDSVSWIRIRPK 936

RESULT 10  
 ID 0933K0 PRELIMINARY; PRT; 1291 AA.  
 AC 0933K0;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Type B cryptic neurotoxin.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=593, and 588;  
 RA Kitama N., Ferreira J.L., Baumstark B.R.;  
 RT "Characterization of six type A strains of Clostridium botulinum that  
 contain type B toxin gene sequences.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF300466; AAI1499.1; -  
 DR EMBL: AF300465; AAI1498.1; -  
 DR InterPro: IPR000395; Bontoxilysin.  
 DR InterPro: IPR006025; Zn\_MTPeptide.  
 DR Pfam: PF01742; Peptidase\_M27; 1.  
 DR PRINTS: PR00760; BONTOTOXILYSIN.  
 DR ProDom: PD001963; Bontoxilysin; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 SQ SEQUENCE 1291 AA; 150843 MW; 7AC1737B0FA5A151 CRC64;

Query Match 31.2% Score 58; DB 2; Length 1291;  
 Best Local Similarity 64.3%; Pred. No. 20;  
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 FNNFTVSFWLRVPR 14  
 I:|:|||||:|:|  
 DB 923 FLDSVSWIRIRPK 936

RESULT 11

O08077  
 ID O08077 PRELIMINARY; PRT; 1291 AA.  
 AC O08077;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE BONT/B.  
 GN BONT/B.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Eklund 17B ATCC25765;  
 RX MEDLINE=94122659; PubMed=7764370;  
 RA Huxon R.A., Collins M.D., East A.K., Thompson D.E.;  
 RT "Nucleotide sequence of the gene coding for non-proteolytic  
 Clostridium botulinum type B neurotoxin: comparison with other  
 Clostridial neurotoxins.";  
 RL Curr. Microbiol. 28:101-110(1994).  
 DR EMBL: X71343; CAA50482.1; -  
 DR HSBP; P10843; 3BTA.  
 DR MEROPS; M27.002; -  
 DR InterPro: IPR000395; Bontoxilysin.  
 DR InterPro: IPR006025; Zn\_MTPeptide.  
 DR Pfam: PF01742; Peptidase\_M27; 1.  
 DR PRINTS: PR00760; BONTOTOXILYSIN.  
 DR ProDom: PD001963; Bontoxilysin; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 SQ SEQUENCE 1291 AA; 150513 MW; 71BCA9E23D69F5AA CRC64;

Query Match 31.2% Score 58; DB 2; Length 1291;  
 Best Local Similarity 64.3%; Pred. No. 20;  
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 FNNFTVSFWLRVPR 14  
 I:|:|||||:|:|  
 DB 923 FLDSVSWIRIRPK 936

RESULT 12  
 ID O08R96 PRELIMINARY; PRT; 1291 AA.  
 AC O08R96;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Neurotoxin.  
 GN BONT/B.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ihara H., Kohda T., Morimoto F., Tsukamoto K., Karasawa T.,  
 RA Nakamura S., Mukamoto M., Kozaki S.;  
 RT "Clostridium botulinum type B neurotoxin associated with infant  
 botulism.";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB084152; BAC22064.1; -  
 SQ SEQUENCE 1291 AA; 150574 MW; 0227CAEFAF58504D CRC64;

Query Match 31.2% Score 58; DB 2; Length 1291;  
 Best Local Similarity 64.3%; Pred. No. 20;  
 Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 FNNFTVSFWLRVPR 14  
 I:|:|||||:|:|  
 DB 923 FLDSVSWIRIRPK 936

```

RESULT 13
O9XAV1 PRELIMINARY: PRT; 430 AA.
ID O9XAV1
AC O9XAV1:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Alkane 1-monooxygenase (EC 1.14.15.1).
GN ALKB.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_Taxid=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CHA0:
RX MEDLINE-21128526; PubMed-11207749;
RA Smith T.H.M.; Koethlißberger M.; Witholt B.; Van Bellem J.B.;
RT "Molecular screening for alkane hydroxylase genes in Gram-negative and
RT Gram-positive bacteria."
RL Environ. Microbiol. 1:307-317(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CHA0:
RA Smith T.H.M.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ009579; CAB51045.2;
DR InterPro: IPR005804; FA_desatL_fam.
DR Pfam: PF00487; FA_desaturase; 1.
KW Monooxygenase; Oxidoreductase.
SQ SEQUENCE 430 AA; 48337 MW; 048E950980783E86 CRC64;

Query Match
Best Local Similarity 30.6%; Score 57; DB 2; Length 430;
Matches 14; Conservative 4; Mismatches 7; Indels 14; Gaps 2;

OY 7 SFMLRVPKY-----SASHLEG-----PSLHWSYGL 31
Db 224 SFMSFLPRVWFSLSAMHLESORLEKGLPLHMKNGV 262

RESULT 14
O9X8T8 PRELIMINARY: PRT; 502 AA.
ID O9X8T8
AC O9X8T8:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative transmembrane protein.
GN SCO3902 OR SH24.24.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_Taxid=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RX MEDLINE-21996410; PubMed-12000953;
RA Bentley S.D.; Chater K.F.; Cerdeno-Tarraga A.-M.; Challis G.L.;
RA Thomson N.R.; James K.D.; Harris D.E.; Quail M.A.; Kleser H.;
RA Harper D.; Bateman A.; Brown S.; Chandra G.; Chen C.W.; Collins M.;
RA Cronin A.; Fraser A.; Goble A.; Hidalgo J.; Hornsby T.; Howarth S.;
RA Huang C.-H.; Kleser T.; Larke L.; Murphy L.; Oliver K.; O'Neill S.;
RA Rabinowitsch E.; Rajandream M.A.; Rutherford K.; Rutter S.;
RA Seeger K.; Saunders D.; Sharp S.; Squares R.; Squares S.; Taylor K.;
RA Warren T.; Wietzorrek A.; Woodward J.; Barrell B.G.; Parkhill J.;
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL: AL939118; CAB42730.1;
KW Transmembrane; Complete proteome.
SQ SEQUENCE 502 AA; 54795 MW; C84F774C0A5AADBC CRC64;

```

```

Query Match
Best Local Similarity 30.6%; Score 57; DB 16; Length 502;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 18 SHLEGPSLHWSYGLRP 33
Db 483 AHPEGPQVHWKGRGP 498

RESULT 15
O90Y63 PRELIMINARY: PRT; 90 AA.
ID O90Y63
AC O90Y63:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Progadoliberin I precursor [contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I
DE (GAP1)]
GN GNRH1 OR GNRH.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_Taxid=8400;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC TISSUE-Forebrain;
RX MEDLINE-21102951; PubMed-11170016;
RA Wang L.; Yoo M.S.; Kang H.M.; Im W.B.; Choi H.S.; Boyerd J.;
RA Kwon H.B.;
RT "Cloning and characterization of cDNAs encoding the GnRH1 and GnRH2
RT precursors from bullfrog (Rana catesbeiana)."
RL J. Exp. Zool. 289:190-201(2001).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: FOREBRAIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT SIGNIFICANTLY HIGHER LEVELS
CC DURING POST-BREEDING. NOT EXPRESSED IN PITUITARY.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL: AF188754; AA05972.1;
DR InterPro: IPR002012; GNRH; 1.
DR InterPro: IPR004079; Gonadoliberin I.
DR Pfam: PF00446; GNRH; 1.
DR PRINTS: PR01541; GONADOLIBRN1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Signal;
KW Pyroglutamate carboxylic acid.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 90 PROGONADOLIBERIN I.
FT PEPTIDE 25 34 GONADOLIBERIN I.
FT PEPTIDE 38 86 GNRH-ASSOCIATED PEPTIDE I
FT FT (BY SIMILARITY).
FT MOD_RES 25 25 PYRROLIDONE CARBOXYLIC ACID
FT FT (BY SIMILARITY).
FT MOD_RES 34 34 AMIDATION (G-35 PROVIDE AMIDE GROUP)
FT FT (BY SIMILARITY).
SQ SEQUENCE 90 AA; 10291 MW; 317203B4E3DA2FE7 CRC64;

Query Match
Best Local Similarity 30.4%; Score 56.5; DB 13; Length 90;
Matches 14; Conservative 4; Mismatches 9; Indels 3; Gaps 2;

OY 5 TWSFWLRVPKVSASH-LEGPSSLHWSYGLRP 33
Db 6 TVVLLAIYVLLSSHHIHGO--HWSYGLRP 33

```

Search completed: September 24, 2003, 17:47:16  
 Job time : 33.964 secs





PD 15-NOV-2001.  
 XX  
 XX 04-MAY-2001; 2001WO-US14363.  
 XX  
 PR 05-MAY-2000; 2000US-202328P.  
 XX  
 PA (APHT-) APHTON CORP.  
 PI Grimes S, Michaeli D, Stevens VC;  
 XX WPI: 2002-049440/06.  
 DR  
 XX  
 PT Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue  
 XX  
 PS Claim 11; Page 8; 43pp: English.  
 XX  
 CC The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX  
 SQ Sequence 28 AA:  
 Query Match 99.3%; Score 147; DB 23; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-15;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QYIRANSKFICITEL--GPSLHWSYGLRP 27  
 |||||||  
 1 QYIRANSKFICITELSGPSLHWSYGLRP 27  
 Db  
 RESULT 2  
 AAU11430  
 ID AAU11430 standard; peptide: 46 AA.  
 XX  
 AC AAU11430;  
 XX  
 DT 12-MAR-2002 (first entry)  
 DE  
 XX Synthetic immunogen peptide 11.  
 DE  
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive;  
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.  
 XX  
 OS Chimeric - Clostridium tetani.  
 OS Chimeric - Mammalia.  
 OS Synthetic.  
 XX  
 FH Key  
 FT Misc-difference 1  
 FT Location/Qualifiers  
 FT /label= OTHER  
 FT /note= "Other- Pyro-glutamic acid or 5-oxo proline"  
 FT 1..10  
 FT /note= "Gonadotropin releasing hormone epitope  
 FT (1..10 aa)"  
 FT 11..16  
 FT /note= "Spacer peptide"  
 FT 17..31  
 FT Peptide

FT /note= "Tetanus toxoid (830-844 aa)"  
 FT 32..37  
 FT /note= "Spacer peptide"  
 FT 38..46  
 FT /note= "Gonadotropin releasing hormone epitope  
 FT (2-10 aa)"  
 FT /note= "Amidated glycine or glycylamide"  
 FT  
 FT Modified-site  
 FT 46  
 FT W0200185763-A2.  
 XX  
 XX 15-NOV-2001.  
 PD  
 XX  
 XX 04-MAY-2001; 2001WO-US14363.  
 XX  
 PR 05-MAY-2000; 2000US-202328P.  
 XX  
 PA (APHT-) APHTON CORP.  
 PI Grimes S, Michaeli D, Stevens VC;  
 XX WPI: 2002-049440/06.  
 DR  
 XX  
 PT Novel synthetic immunogen for inducing immune response against  
 PT gonadotropin releasing hormone, comprises fusion peptide having  
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
 PT epitope or its analogue  
 XX  
 PS Claim 11; Page 12; 43pp: English.  
 XX  
 CC The invention relates to a synthetic immunogen for inducing specific  
 CC antibodies against gonadotropin releasing hormone (GnRH) also known  
 CC as luteinising hormone releasing hormone, LHRH) comprising a fusion  
 CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
 CC immunomimic peptide epitope or its analogue. The synthetic  
 CC immunogen is useful inducing an immune response against GnRH in an  
 CC animal subject, and as such is useful as a contraceptive and in the  
 CC treatment of diseases such as cancer (of the breast, uterus and other  
 CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic  
 CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
 CC high and specific anti-GnRH antibody titres. The present sequence  
 CC is a synthetic immunogen of the invention.  
 XX  
 SQ Sequence 46 AA:  
 Query Match 91.9%; Score 136; DB 23; Length 46;  
 Best Local Similarity 93.1%; Pred. No. 1.3e-13;  
 Matches 27; Conservative 0; Mismatches 0; Indels 2; Gaps 1;  
 QY 1 QYIRANSKFICITEL--GPSLHWSYGLRP 27  
 |||||||  
 17 QYIRANSKFICITELSGPSLHWSYGLRP 45  
 Db  
 RESULT 3  
 AAR62723  
 ID AAR62723 standard; peptide: 47 AA.  
 XX  
 AC AAR62723;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 17-SEP-1995 (first entry)  
 DE  
 XX LHRH-containing immunogenic peptide.  
 XX  
 KW Helper T cell epitope; universal immune stimulator; invasins; haptens;  
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;  
 KW androgen-dependent carcinoma; antitumour; infertility;  
 KW tetanus toxin.  
 XX  
 OS Synthetic.  
 OS  
 FH Key  
 FT Location/Qualifiers

```

FT Domain 1..16 /note= "invasin domain"
FT Domain 19..35 /note= "tetanus toxin helper T cell epitope"
FT Domain 38..47 /note= "LHRH hapten"
XX
XX MO9425060-A1.
XX
XX 10-NOV-1994.
XX
XX 28-APR-1994; 94WO-US04832.
XX
XX 27-APR-1993; 93US-0057166.
XX 14-APR-1994; 94US-0229275.
XX
XX (LADD/) LADD A E.
XX (WANG/) WANG C Y.
XX (ZAMB/) ZAMB T.
XX
XX Ladd AE, Wang CY, Zamb T;
XX
XX WPI: 1994-357910/44.
XX
XX Immunogenic luteinising hormone releasing hormone peptide(s) -
XX PT that suppress LHRH activity in males and females
XX
XX Claim 8; Page 88; 213pp: English.
XX
XX Synthetic immunogenic peptides are provided in which a universal immune
XX stimulator is linked to a peptide or protein hapten containing B cell
XX and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX potent immune responses to the coupled peptide or protein. The
XX stimulator consists of (A) a promiscuous helper T cell epitope (Th)
XX which elicits an immune response to the coupled peptide in members of
XX a heterogeneous population expressing diverse HLA phenotypes, and (B)
XX an adjuvant peptide sequence from the invasins protein of Yersinia.
XX Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
XX invasins and Th domains and between the immune stimulator and hapten
XX components. When the hapten is LHRH, then optionally the invasins domain
XX can be omitted from the immune stimulator component.
XX CC The present sequence represents an LHRH-containing immunogenic peptide
XX as above which can be used as a potent vaccine for treating e.g.
XX prostatic hyperplasia, androgen-dependent carcinoma, prostatic
XX carcinoma, testicular carcinoma, endometriosis, benign uterine tumours,
XX recurrent functional ovarian cysts, (severe) premenstrual syndrome or
XX oestrogen-dependent breast cancer, or for induction of infertility.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 47 AA;
XX
XX Query Match 80.7%; Score 119.5; DB 15; Length 47;
XX Best Local Similarity 88.9%; Pred. No. 4.7e-11;
XX Matches 24; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
XX
XX 1 OYIKANSKFIGITELPSLHWSYGLRP 27
XX ||||||||||||| |||||||
XX 21 OYIKANSKFIGITELGGE-HWSYGLRP 46
XX
XX RESULT 4
XX AAR62701
XX ID AAR62701 standard; peptide; 27 AA.
XX
XX AC AAR62701;
XX
XX 25-MAR-2003 (updated)
XX DT 10-SEP-1995 (first entry)
XX
XX LHRH-containing immunogenic peptide.
XX
XX Helper T cell epitope; universal immune stimulator; invasins; hapten;
XX vaccine; LHRH; luteinising hormone releasing hormone; prostate;
XX

```

```

KW androgen-dependent carcinoma; antitumour; infertility;
KW tetanus toxin.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FT Domain 1..17 /note= "tetanus toxin helper T cell epitope"
XX FT Domain 18..27 /note= "LHRH hapten"
XX
XX MO9425060-A1.
XX
XX 10-NOV-1994.
XX
XX 28-APR-1994; 94WO-US04832.
XX
XX 27-APR-1993; 93US-0057166.
XX 14-APR-1994; 94US-0229275.
XX
XX (LADD/) LADD A E.
XX (WANG/) WANG C Y.
XX (ZAMB/) ZAMB T.
XX
XX Ladd AE, Wang CY, Zamb T;
XX
XX WPI: 1994-357910/44.
XX
XX Immunogenic luteinising hormone releasing hormone peptide(s) -
XX PT that suppress LHRH activity in males and females
XX
XX Claims 8, 12; Page 84; 213pp: English.
XX
XX Synthetic immunogenic peptides are provided in which a universal immune
XX stimulator is linked to a peptide or protein hapten containing B cell
XX and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX potent immune responses to the coupled peptide or protein. The
XX stimulator consists of (A) a promiscuous helper T cell epitope (Th)
XX which elicits an immune response to the coupled peptide in members of
XX a heterogeneous population expressing diverse HLA phenotypes, and (B)
XX an adjuvant peptide sequence from the invasins protein of Yersinia.
XX Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
XX invasins and Th domains and between the immune stimulator and hapten
XX components. When the hapten is LHRH, then optionally the invasins domain
XX can be omitted from the immune stimulator component.
XX CC The present sequence represents an LHRH-containing, invasins-free
XX immunogenic peptide as above which can be used as a potent vaccine for
XX treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
XX prostatic carcinoma, testicular carcinoma, endometriosis, benign
XX uterine tumours, recurrent functional ovarian cysts, (severe)
XX premenstrual syndrome or oestrogen-dependent breast cancer, or for
XX induction of infertility.
XX CC This sequence is particularly preferred.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 27 AA;
XX
XX Query Match 76.0%; Score 112.5; DB 15; Length 27;
XX Best Local Similarity 85.2%; Pred. No. 3.1e-10;
XX Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
XX
XX 1 OYIKANSKFIGITELPSLHWSYGLRP 27
XX ||||||||||||| |||||||
XX 3 OYIKANSKFIGITELF--HWSYGLRP 26
XX
XX RESULT 5
XX AAY92639
XX ID AAY92639 standard; protein; 750 AA.
XX
XX AC AAY92639;
XX
XX 10-AUG-2000 (first entry)
XX

```

```

XX XX Mutant human prostate specific membrane antigen construct, hPSM5.1.
DE XX
XX XX Prostate specific membrane antigen; immunogenized construct; mutant;
KW KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX XX
OS Homo sapiens.
OS Synthetic.
XX XX
FH Key Location/Qualifiers
FT Peptide 21..41
FT /label= P30
FT /note= "foreign epitope"
FT Peptide 305..319
FT /label= P2
FT /note= "foreign epitope"
XX XX
XX XX WO200020027-A2.
XX XX
XX XX 13-APR-2000.
XX XX
XX XX 05-OCT-1999; 99WO-DK00525.
XX XX
XX XX 05-OCT-1998; 98DK-0001261.
XX XX 20-OCT-1998; 98US-0105011.
XX XX
XX XX (MEBI-) M & E BIOTECH AS.
XX XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX XX
XX XX WPI; 2000-349917/30.
XX XX
PT Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX XX
XX XX Example 1; Page -: 220pp; English.
XX XX
XX XX AAY92627-49 are mutant immunogenized human prostate specific membrane
XX antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
XX P30). The immunogenic analogues of PSM can be used in the claimed method
XX as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX antibody binding regions and cysteine residues involved in disulfide
XX bonds are preserved in the immunogenized forms. The method is used for
XX inducing immune responses against weakly immunogenic cell-associated
XX peptide antigens (PA) such as those associated with cancers
XX (self-proteins), e.g. human prostate specific membrane antigen (PSM),
XX heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX method comprises effecting simultaneous presentation by antigen producing
XX cells (APCs) of the animals immune system of: (1) at least 1 CTL
XX (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
XX B-cell group derived from the cell-associated PA; and (2) at least 1
XX first T helper cell group which is foreign to the animal. Analogues of
XX human PSM, human Her2 and human/murine FGF8b comprising a substantial
XX part of all known and predicted CTL and B-cell epitopes of the respective
XX PA and including at least one foreign T helper epitope are also claimed.
XX The method is used to treat prostate, prostate/breast or breast cancer
XX when the PA is human PSM, FGF8b and Her2, respectively.
XX Note: This sequence was constructed from the wild type human PSM
XX (AAY92619), which appears on pages 184-187 of the specification.
XX XX
SQ Sequence 750 AA;

```

```

Query Match 58.1%; Score 86; DB 21; Length 750;
Best Local Similarity 70.4%; Pred. No. 0.00016;
Matches 19; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 QYIKANSKFIGITELGSLHWSYGLRP 27
DB 305 QYIKANSKFIGITELRSLKVPYNWGP 331

```

```

RESULT 6
AAM81329
ID AAM81329 standard; Protein; 158 AA.
XX XX
XX XX AAM81329;
AC XX
XX XX 21-APR-1999 (first entry)
DE XX
XX XX TNF2-4, a TNF-alpha analogue.
XX XX
KW Human tumour necrosis factor-alpha; TNF-alpha; TNF-alpha analogue;
KW vaccine; rheumatoid arthritis; Crohn's disease; ulcerative colitis;
KW cancer; disseminated sclerosis; diabetes; psoriasis; osteoporosis;
KW asthma.
XX XX
OS Synthetic.
OS Homo sapiens.
XX XX
XX XX WO9846642-A1.
XX XX
XX XX 22-OCT-1998.
XX XX
XX XX 15-APR-1998; 98WO-DK00157.
XX XX
XX XX 24-APR-1997; 97US-0044187.
XX XX 15-APR-1997; 97DK-0000418.
XX XX
XX XX (FERR ) FARM LAB FERRING AS.
XX XX
PI Dalum I, Elsnor H, Jensen MR, Mouritsen S;
PI N-PSDB; AAV68418.
XX XX
XX XX WPI; 1998-594561/50.
XX XX
XX XX N-PSDB; AAV68418.
XX XX

```

```

PT Modified human tumour necrosis factor-alpha - comprises
PT immunodominant T cell epitope, useful in vaccines to treat or
PT prevent TNF-associated diseases, e.g. cancer
XX XX
XX XX Example 1; Page 69-70; 134pp; English.
XX XX

```

```

XX XX The present sequence represents a modified human tumour necrosis
XX factor-alpha (TNF-alpha) analogue. The analogues have no residual
XX TNF activity and are immunogenic in a large proportion of the human
XX population (by using promiscuous epitopes). The TNF-alpha analogue
XX is able to generate, in humans, neutralizing antibodies to wild-type
XX human TNF alpha, has at least one fragment of TNF substituted by a
XX peptide containing an immunodominant T-cell epitope, and at least one
XX TNF-alpha B-cell epitope. The substitution causes a significant change
XX in the amino acid sequence of any one of the strands in the front
XX beta-sheet, any of the connecting loops or any of the B', I or D strands
XX in the back beta-sheet. The TNF-alpha analogues are used as vaccines for
XX treatment or prevention of diseases associated with excessive release
XX or activity of TNF-alpha, e.g. rheumatoid arthritis, Crohn's disease,
XX ulcerative colitis, cancer of any sort, disseminated sclerosis, diabetes,
XX psoriasis, osteoporosis and asthma.
XX XX

```

```

SQ Sequence 158 AA;
Query Match 55.4%; Score 82; DB 19; Length 158;
Best Local Similarity 89.5%; Pred. No. 0.00011;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 QYIKANSKFIGITELGSL 19
DB 115 QYIKANSKFIGITELGRL 133

```

```

RESULT 7
ABB07281
ID ABB07281 standard; Protein; 158 AA.
XX XX
XX AC ABB07281;
XX XX

```



```

DT 10-AUG-2000 (first entry)
XX
XX Mutant human prostate specific membrane antigen construct, hPSM10.1.
DE Prostate specific membrane antigen; immunogenized construct; mutant;
XX vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Peptide 21..41
XX /label= P30
XX /note= "foreign epitope"
XX Peptide 674..688
XX /label= P2
XX /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
XX 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX Gautam A, Birk P, Karlsson G;
XX
XX WPI: 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
XX peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page -: 220pp; English.
XX
XX AAY92627-49 are mutant immunogenized human prostate specific membrane
XX antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
XX P30). The immunogenic analogues of PSM can be used in the claimed method
XX as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX antibody binding regions and cysteine residues involved in disulfide
XX bonds are preserved in the immunogenized forms. The method is used for
XX inducing immune responses against weakly immunogenic cell-associated
XX peptide antigens (PA) such as those associated with cancers
XX (self-proteins), e.g. human prostate specific membrane antigen (PSM),
XX heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX method comprises effecting simultaneous presentation by antigen producing
XX cells (APCs) of the animals immune system of: (1) at least 1 CTL
XX (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
XX B-cell group derived from the cell-associated PA; and (2) at least 1
XX first T helper cell group which is foreign to the animal. Analogues of
XX human PSM, human Her2 and human/murine FGF8b comprising a substantial
XX part of all known and predicted CTL and B-cell epitopes of the respective
XX PA and including at least one foreign T helper epitope are also claimed.
XX The method is used to treat prostate, prostate/breast or breast cancer
XX when the PA is human PSM, FGF8b and Her2, respectively.
XX Note: This sequence was constructed from the wild type human PSM
XX (AAY92619), which appears on pages 184-187 of the specification.
XX
XX Sequence 750 AA:
XX
XX Query Match 55.4%; Score 82; DB 21; Length 750;
XX Best Local Similarity 70.4%; Pred. No. 0.00066;
XX Matches 19; Conservative 0; Mismatches 4; Indels 4; Gaps 1;
XX
XX 1 QYIKANSKFIGITEL---GPSLHWSTY 23
XX |||||
XX 674 QYIKANSKFIGITELHVIAPSSHNKY 700

```

```

RESULT 10
ID AAY92641
XX AAY92641 standard; Protein; 750 AA.
XX
XX AAY92641;
XX
XX 10-AUG-2000 (first entry)
XX
XX Mutant human prostate specific membrane antigen construct, hPSM10.0.
XX
XX Prostate specific membrane antigen; immunogenized construct; mutant;
XX vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Peptide 674..688
XX /label= P2
XX /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
XX 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX Gautam A, Birk P, Karlsson G;
XX
XX WPI: 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
XX peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page -: 220pp; English.
XX
XX AAY92627-49 are mutant immunogenized human prostate specific membrane
XX antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
XX P30). The immunogenic analogues of PSM can be used in the claimed method
XX as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX antibody binding regions and cysteine residues involved in disulfide
XX bonds are preserved in the immunogenized forms. The method is used for
XX inducing immune responses against weakly immunogenic cell-associated
XX peptide antigens (PA) such as those associated with cancers
XX (self-proteins), e.g. human prostate specific membrane antigen (PSM),
XX heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX method comprises effecting simultaneous presentation by antigen producing
XX cells (APCs) of the animals immune system of: (1) at least 1 CTL
XX (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
XX B-cell group derived from the cell-associated PA; and (2) at least 1
XX first T helper cell group which is foreign to the animal. Analogues of
XX human PSM, human Her2 and human/murine FGF8b comprising a substantial
XX part of all known and predicted CTL and B-cell epitopes of the respective
XX PA and including at least one foreign T helper epitope are also claimed.
XX The method is used to treat prostate, prostate/breast or breast cancer
XX when the PA is human PSM, FGF8b and Her2, respectively.
XX Note: This sequence was constructed from the wild type human PSM
XX (AAY92619), which appears on pages 184-187 of the specification.
XX
XX Sequence 750 AA:
XX
XX Query Match 55.4%; Score 82; DB 21; Length 750;
XX Best Local Similarity 70.4%; Pred. No. 0.00066;
XX Matches 19; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

```



CC sequence from the present invention.

XX Sequence 31 AA;

Query Match 55.1%; Score 81.5; DB 21; Length 31;  
Best Local Similarity 61.3%; Pred. No. 2.2e-05;  
Matches 19; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

QY 1 OYKANSKFIGITELG-----PSLHWS 22  
1 OYKANSKFIGITELGHEIKKVLVPGCHGS 31

Db

RESULT 13  
AAU11423  
ID AAU11423 standard; peptide: 33 AA.

XX AAU11423;

DT 12-MAR-2002 (first entry)

XX Synthetic immunogen peptide 4.

DE Gonadotrophin releasing hormone; GnRH; synthetic immunogen;  
KW luteinizing hormone releasing hormone; LHRH; contraceptive;  
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;  
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;  
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX OS Chimeric - Plasmodium falciparum.

OS OS Chimeric - Mammalia.  
OS Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..20 /note= "Malaria CSP protein (378-398 aa)"

FT Peptide 21..24 /note= "Spacer peptide"

FT Peptide 25..33 /note= "Gonadotrophin releasing hormone epitope"

FT Modified-site 33 /note= "Amidated glycine or glycylamide"

XX W0200185763-A2.

XX 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US14363.

XX 05-MAY-2000; 2000US-202328P.

XX (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;

XX WPI; 2002-049440/06.

XX Novel synthetic immunogen for inducing immune response against  
PT gonadotrophin releasing hormone, comprises fusion peptide having  
PT promiscuous helper T-cell peptide epitope and immunomimic peptide  
PT epitope or its analogue

XX Claim 11; Page 8; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific  
CC antibodies against gonadotrophin releasing hormone (GnRH) also known  
CC as luteinizing hormone releasing hormone, LHRH) comprising a fusion  
CC peptide which comprises a promiscuous helper T-cell peptide epitope and  
CC immunomimic peptide epitope or its analogue. The synthetic  
CC immunogen is useful inducing an immune response against GnRH in an  
CC animal subject, and as such is useful as a contraceptive and in the  
CC treatment of diseases such as cancer (of the breast, uterus and other  
CC gynaecological cancer), endometriosis, uterine fibroids, benign prostatic

CC hypertrophy and prostate cancer. The immunogen is effective in eliciting  
CC high and specific anti-GnRH antibody titres. The present sequence  
CC is a synthetic immunogen of the invention.

XX Sequence 33 AA;

Query Match 55.1%; Score 81.5; DB 23; Length 33;  
Best Local Similarity 66.7%; Pred. No. 2.4e-05;  
Matches 16; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 4 KANSKFIGITELGPSLHWSYGLRP 27  
10 KASSVF-NVNSGPSLHWSYGLRP 32

Db

RESULT 14  
AAR83561  
ID AAR83561 standard; peptide: 29 AA.

XX AAR83561;

DT 13-JUN-1996 (first entry)

XX Ige CH4 region contg. peptide immunogen for treating allergies.

XX Ige; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;  
KW vaccine; allergy; antibody; constant heavy chain.

XX Synthetic.

XX W09526365-A1.

XX 05-OCT-1995.

XX 24-MAR-1995; 95WO-US03741.

XX 25-OCT-1994; 94US-0328912.

XX 28-MAR-1994; 94US-0218461.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 1995-351297/45.

XX Synthetic peptide-based immunogen contg. Ige CH4 peptide and helper  
PT T cell epitope - useful for eliciting antibody prodn. for allergy  
PT treatment

XX Claim 5; Page 68-69; 87pp; English.

XX AAR82592-R82600 and AAR83560-R83581 are peptide immunogens that are  
CC useful in vaccines for treating allergic reactions. In the immunogens,  
CC an Ige CH4 peptide is attached C-terminally to a series of amino acids  
CC including a helper T cell epitope. The immunogen may also opt. contain  
CC a fatty acid or fatty acid derivative, an invasion domain or alpha-NH2.  
CC The immunogen produces high titres of antibodies to the effector site  
CC in human Ige heavy chain (the CH4 domain peptide) which inhibit mast  
CC cell activation and reduce allergen-induced Ige prodn. The immunogens  
CC may be used in either a radially branching multimeric form or a  
CC linearly arranged monomeric form.

XX Sequence 29 AA;

Query Match 54.1%; Score 80; DB 16; Length 29;  
Best Local Similarity 100.0%; Pred. No. 3.5e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OYKANSKFIGITELG 16  
3 OYKANSKFIGITELG 18

Db

Search completed: September 24, 2003, 17:40:04  
 Job time : 32.6454 secs

```

RESULT 15
AAV82636
ID AAV82636 strand: peptide; 32 AA.
XX
XX
AC AAV82636:
XX
XX
DT 07-AUG-2000 (first entry)
XX
XX
DE Tetanus toxoid T cell epitope and Der pII B cell epitope peptide.
XX
XX
KW T cell epitope; B cell epitope; allergy; allergen; antigenic;
KW antiallergic; antiasthmatic; antinflammatory; dermatological;
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KW atopic dermatitis; acute urticaria; chronic urticaria;
KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
KW anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX
XX
OS Dermatophagoides pteronyssinus.
OS Clostridium tetani.
OS Synthetic.
XX
XX
PN WO200006694-A2.
XX
XX
PD 10-FEB-2000.
XX
XX
PF 20-JUL-1999; 99WO-BE00092.
XX
XX
PR 30-JUL-1998; 98EP-0870167.
XX
XX
PA (UNIO ) UCB SA.
XX
XX
PI Saint-Remy J, Jacquemin M;
XX
XX
DR WPI; 2000-422470/36.
XX
XX
PT New compound for prevention and treatment of allergies comprises at
PT least one allergen antigenic determinant recognized by a B cell and at
PT least one antigenic determinant which does not trigger T cell
PT activation -
XX
XX
PS Claim 8; Page 35; 50pp; English.
XX
XX
CC The present invention describes a compound (I) for the prevention and/or
CC treatment of allergy. The compound comprises at least one allergen
CC antigenic determinant (I) recognised by a B cell or an antibody secreted
CC by a B cell of a non-atopic individual and at least one antigenic
CC determinant (II) different from the allergen that triggers T cell
CC activation. (I) has antiallergic, antiasthmatic, antinflammatory,
CC dermatological and immunosuppressive activities, and can be used in a
CC vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to
CC treat and/or prevent allergies or a disease of allergic origin,
CC especially hypersensitivities. These include rhinitis, sinusitis,
CC bronchial asthma, atopic dermatitis, some forms of acute and chronic
CC urticaria, gastro-intestinal syndromes associated with the ingestion of
CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions
CC associated with drug hypersensitivities and/or a mixture of these. The
CC use of (I) in the treatment of allergic conditions avoids the need for
CC drug treatment, which often causes undesirable side-effects. Also, prior
CC art drug therapies alleviate symptoms, but do not influence their
CC causes, however (I) actually combats the cause of an allergic reaction.
CC The present sequence represents a specifically claimed compound peptide
CC sequence from the present invention.
XX
XX
SQ Sequence 32 AA:

```

```

Query Match          54.1%; Score 80; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QYIKANSKFGITEIG 16
   ||||||||||||||||
DB 1 QYIKANSKFGITEIG 16

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 16:46:20 ; Search time 8.29917 Seconds  
(without alignments)  
142.750 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYIKANSKRTIGTELGPSTLHWSYGLRPX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PC105.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfill1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query	Match	Length	ID	Description
	1	119.5	80.7	47	1	US-08-446-692-35	Sequence 35, Appl
	2	119.5	80.7	47	2	US-08-488-351A-35	Sequence 35, Appl
	3	112.5	76.0	27	1	US-08-446-692-13	Sequence 13, Appl
	4	112.5	76.0	27	2	US-08-488-351A-13	Sequence 13, Appl
	5	80.7	54.1	37	1	US-08-446-692-57	Sequence 57, Appl
	6	80.7	54.1	37	1	US-08-446-692-63	Sequence 63, Appl
	7	80.7	54.1	37	2	US-08-488-351A-57	Sequence 57, Appl
	8	80.7	54.1	37	2	US-08-488-351A-63	Sequence 63, Appl
	9	77.7	52.0	24	4	US-08-464-496-25	Sequence 25, Appl
	10	77.7	52.0	24	4	US-08-464-496-25	Sequence 25, Appl
	11	77.7	52.0	24	4	US-08-197-484-110	Sequence 110, Appl
	12	77.7	52.0	24	4	US-08-197-484-110	Sequence 110, Appl
	13	77.7	52.0	24	5	PCT-US93-02121-110	Sequence 110, Appl
	14	77.7	52.0	27	4	US-08-464-496-32	Sequence 32, Appl
	15	77.7	52.0	27	5	PCT-US92-07218-32	Sequence 32, Appl
	16	76.5	51.4	31	4	US-08-432-483A-17	Sequence 2, Appl
	17	76.5	51.4	31	4	US-08-945-289-2	Sequence 2, Appl
	18	74.4	50.0	15	2	US-08-319-704-10	Sequence 10, Appl
	19	74.4	50.0	15	2	US-08-661-052-6	Sequence 6, Appl
	20	74.4	50.0	15	2	US-08-460-502-7	Sequence 7, Appl
	21	74.4	50.0	15	3	US-09-046-373-2	Sequence 2, Appl
	22	74.4	50.0	15	3	US-09-188-082-6	Sequence 6, Appl
	23	74.4	50.0	15	4	US-09-364-088-6	Sequence 6, Appl
	24	74.4	50.0	15	4	US-09-102-716-6	Sequence 6, Appl
	25	74.4	50.0	15	4	US-09-148-711A-7	Sequence 7, Appl
	26	74.4	50.0	15	4	US-08-716-249-4	Sequence 4, Appl
	27	74.4	50.0	15	5	PCT-US93-11703-69	Sequence 69, Appl

28	74	50.0	16	3	US-09-248-588-55	Sequence 55, Appl
29	74	50.0	17	2	US-08-446-692-4	Sequence 4, Appl
30	74	50.0	17	2	US-08-488-351A-4	Sequence 4, Appl
31	74	50.0	17	3	US-09-100-409A-40	Sequence 40, Appl
32	74	50.0	17	4	US-08-472-701-23	Sequence 23, Appl
33	74	50.0	17	5	PCT-US95-08596-23	Sequence 23, Appl
34	74	50.0	17	5	PCT-US95-13841-7	Sequence 7, Appl
35	74	50.0	31	5	PCT-US93-11703-63	Sequence 63, Appl
36	74	50.0	853	4	US-08-913-880C-17	Sequence 17, Appl
37	74	50.0	858	4	US-08-913-880C-16	Sequence 16, Appl
38	74	50.0	860	4	US-08-913-880C-15	Sequence 15, Appl
39	74	50.0	862	4	US-08-913-880C-14	Sequence 14, Appl
40	74	50.0	862	4	US-08-913-880C-13	Sequence 13, Appl
41	74	50.0	866	4	US-08-913-880C-12	Sequence 12, Appl
42	74	50.0	874	4	US-08-913-880C-11	Sequence 11, Appl
43	74	50.0	875	4	US-08-913-880C-10	Sequence 10, Appl
44	74	50.0	1315	4	US-08-913-880C-1	Sequence 1, Appl
45	73	49.3	29	3	US-09-075-257A-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-08-446-692-35  
; Sequence 35, Application US/08446692  
; Patent No. 5759551  
; GENERAL INFORMATION:  
; APPLICANT: Ladd, Anna  
; APPLICANT: Wang, Chang Yi  
; APPLICANT: Zamb, Timothy  
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maria C.H. Lin  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: US  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,692  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria C.H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)415-8745  
; TELEFAX: (516)751-6849  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 47 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-446-692-35

Query Match 80.7%; Score 119.5; DB 1; Length 47;  
Best local similarity 88.9%; Pred. No. 7.8e-12;  
Matches 24; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 QYIKANSKRTIGTELGPSTLHWSYGLRP 27  
|||||  
DB 21 QYIKANSKRTIGTELGPSTLHWSYGLRP 46

RESULT 2  
US-08-488-351A-35  
Sequence 35, Application US/08488351A  
Patent No. 5843446  
GENERAL INFORMATION:  
APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang Yi  
APPLICANT: Zamb, Timothy  
TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maria C.H. Lin  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: US  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,351A  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/446,692  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/229,275  
FILING DATE: 14-APR-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/057,166  
FILING DATE: 27-APR-1992  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria C.H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4146 US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)415-8745  
TELEFAX: (516)751-6849  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-488-351A-35

Query Match 80.7%; Score 119.5; DB 2; Length 47;  
Best Local Similarity 88.9%; Pred. No. 7.8e-12;  
Matches 24; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 QYKANSKFIGITELGPSLHWSYGLRP 27  
DB 21 QYKANSKFIGITELGGE-HWSYGLRP 46

RESULT 3  
US-08-446-692-13  
Sequence 13, Application US/08446692  
Patent No. 5759551  
GENERAL INFORMATION:  
APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang Yi  
APPLICANT: Zamb, Timothy  
TITLE OF INVENTION: Immunogenic LHRH peptide constructs

TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maria C.H. Lin  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: US  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,692  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria C.H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4146 US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)415-8745  
TELEFAX: (516)751-6849  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-446-692-13

Query Match 76.0%; Score 112.5; DB 1; Length 27;  
Best Local Similarity 85.2%; Pred. No. 5.2e-11;  
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 QYKANSKFIGITELGPSLHWSYGLRP 27  
DB 3 QYKANSKFIGITELG--HWSYGLRP 26

RESULT 4  
US-08-488-351A-13  
Sequence 13, Application US/08488351A  
Patent No. 5843446  
GENERAL INFORMATION:  
APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang Yi  
APPLICANT: Zamb, Timothy  
TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maria C.H. Lin  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: US  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,351A  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/446,692  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424